ABSTRACT

Title of Dissertation:	ENZYME-FREE UBIQUITIN LIGATION. FROM NATIVE CHEMICAL LIGATION AND SPIN LABELING TO DIMERIZATION BY INTER-UBIQUITIN MIMICKING LINKAGES AND PH-DEPENDENT CONFORMATIONAL SWITCH
	Karina Herrera-Guzman, Doctor of Philosophy, 2016
Dissertation directed by:	Dr. David Fushman, Department of Chemistry and Biochemistry

The delicate balance between the production and disposal of proteins is vital for the changes required in the cell to respond to various stimuli. Ubiquitination is a protein modification with a range of signaling outcomes when ubiquitin is attached to a protein through a highly ordered enzymatic cascade process. Understanding ubiquitination is a growing field and nowadays the application of chemical reactions allows the isolation of quantitative materials for structural studies. Therefore, in this dissertation are described and explored some of these suitable chemical methodologies to achieve the polymerization of ubiquitin while bypassing the enzymatic control, with the purpose of examining if these chemical modifications have a direct impact in the structure of ubiquitin. First, the possibility of incorporating non-natural lysine analogs known as mercaptolysines into the polypeptide chain of

ubiquitin was explored when they were attached to ubiquitin by native chemical ligation at its C terminus. The sulfhydryl group was used for the attachment of a paramagnetic label to map the surface of ubiquitin. Second, the condensation mediated by silver nitrate was used for the dimer assembly. In particular, the main focus was on examining whether various schemes for orthogonal protection and deprotection of each monomer have an impact on the reaction yield, since the synthetic strategy has been previously attempted successfully. Third, the formation of ubiquitin dimers was approached by building an inter-ubiquitin linkage mimicking the isopeptide bond with two approaches, the classic disulfide exchange as well as the thiol-ene click reaction by thermal initiation in aqueous conditions. After assembling the dimeric units, they were studied by Nuclear Magnetic Resonance, in order to establish the conformational state of the resulting di-ubiquitins and whether it depends on the pH conditions. The latter is a very important concept since some ligands have a preferred affinity when the protein-protein hydrophobic patches are in close proximity.

ENZYME-FREE UBIQUITIN LIGATION. FROM NATIVE CHEMICAL
LIGATION AND SPIN LABELING TO DIMERIZATION BY INTER-UBIQUITIN
MIMICKING LINKAGE AND PH-DEPENDENT CONFORMATIONAL SWITCH

by

Karina Herrera-Guzman

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor in Philosophy
2016

Advisory Committee:
Professor Dr. David Fushman, Chair
Dr. Jeffrey Davis
Dr. Douglas Julin
Dr. Daniel Falvey
Dr. Johnathan Dinman

Acknowledgements

I would like to thank first to my advisor Dr. David Fushman for his support through my PhD studies not only as a professor but also as a good guidance to keep going and work hard to achieve any goal I traced to give shape to my research. The most valuable moments for me are the long sessions discussing concepts which helped me to develop an outside of the box way of thinking on how proteins behave and how to jump the bridge from simple organic chemistry synthesis to more complex systems. In addition, I am thankful to him for not leaving me alone in my goal of completing my graduate studies, his confidence on me developing the new skills in the laboratory but most importantly his patience and time sharing with me his knowledge in NMR. In a personal matter, I also want to thank him for all his support in the difficult times I experienced and his good character.

Second, I would like to thank the members of the advisory committee for their comments and guidance in this dissertation.

Third, to Dr. Daoning Zhang and Dr. Yue Li for their help with NMR and MS, respectively.

Fourth, I would like to thank all the members of the Fushman lab because without them the work in the laboratory wouldn't be so enjoyable; all of them have a very special place in my life. Apurva, Tanuja and Urszula for being my first friends in which I could find help and understanding; Adithya and Dulith for the good times; Carlos, Mark and Rajesh for their help with developing ideas; Christina for her patience when her name is confused with mine as well as the good talks about food

Table of Contents

List of Tables

List of Figures

List of Schemes

List of Abbreviations

3D: Three-dimension

ATP: Adenosine triphosphate

CBD: Chitin binding domain

CD: Circular dichroism

CMA: Chaperone mediated autophagy

CSP: Chemical shift perturbation

DEAE: Diethylaminoethyl cellulose

DMF: N,N-dimethyl formamide

DMSO: Dimethyl sulfoxide

DNA: Deoxyribonucleic acid

DTT: Dithiothreitol

DUB: Deubiquitinating enzyme

ESI-MS: Electrospray ionization – Mass spectrometry

EWG: Electrowithdrawing group

GFP: Green fluorescent protein

HSQC: Heteronuclear single quantum coherence spectroscopy

LG: Leaving group

MesNa: 2-Mercaptoethanesulfonic acid, sodium salt

MPAA: 4-Mercaptophenylacetic acid

MTSL: Methanesulfonothioate spin label

NCL: Native Chemical Ligation

NMR: Nuclear Magnetic Resonance

PRE: Paramagnetic relaxation enchancement

RDS: Rate Determining Step

RNA: Ribonucleic acid

SDS: Sodium dodecyl sulfate

SDS-PAGE: Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SPPS: Solid Phase Peptide Synthesis

TFA: Trifluoroacetic acid

THF: Tetrahydrofuran

TOCSY: Total correlation spectroscopy

Ub: Ubiquitin

UBD: Ubiquitin Binding Domain

Ubl: Ubiquitin-like

UPS: Ubiquitin proteasome system

UV: Ultraviolet

WT: Wild type

Introduction

In nature, cells have a system of post-translational modification of proteins to achieve and address signals in almost every pathway. Regulated protein degradation is essential for cell survival because it is necessary to respond immediately to environmental changes in order to keep homeostasis.[1] In the cell a polypeptide chain must have a proper three-dimensional fold in order to function properly. A proper folding, however, is not a simple task because the correct amino acid sequence is not the only factor, and as a consequence misfolding of proteins is an inevitable process. Three protein quality control systems efficiently remove those misfolded proteins from various cellular compartments (cytoplasm, nucleus, endoplasmic reticulum) 1) Ubiquitin proteasome system (UPS), 2) chaperone mediated autophagy (CMA) and 3) macroautophagy.[2] Soluble misfolded proteins are degraded in the UPS once they are tagged by ubiquitin (Ub). Those misfolded proteins which expose the sequence KFERQ are degraded by the CMA once they are recognized by the chaperon heat-shock cognate 70 (Hsc70) and those proteins which persist despite the control of the UPS and CMA form aggregates and degraded by macroautophagy (Figure 1).[3]
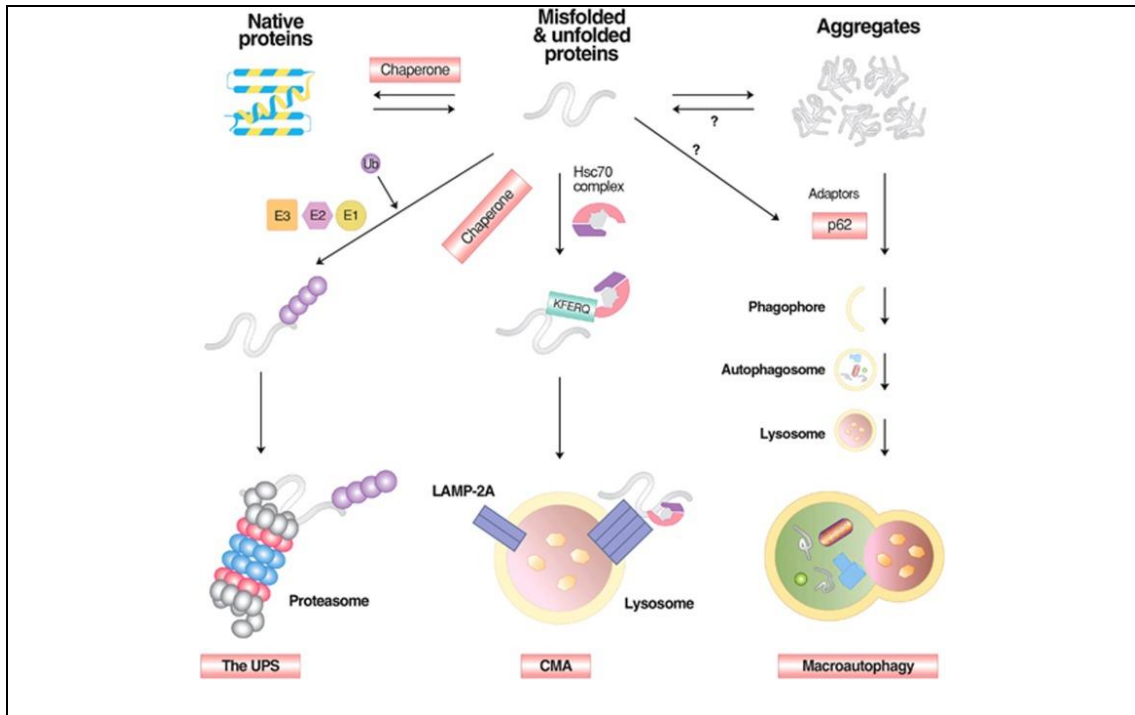
Figure 1. The degradation of misfolded proteins by varios cellular proteolytic pathways.[4]

However, some polypeptides tend to misfold spontaneously and aggregate into oligomers enriched in β-sheet content.[2] In neurons, for instance, the protein quality control is *per se* challenging due to their unique structure in terms on how the dendrites and axons expand. Genetic mutations in proteins such as huntingtin in Huntington's disease,[5] α-synuclein in Parkinson's,[6] and other neurodegenerative diseases increase the formation of β-sheet aggregates. Moreover, these aggregates could become cytotoxic if they are not degraded. The task of degradation then becomes even more puzzling if the components of the UPS, CMA and macroautophagy are downregulated in terms of expression and activity due to aging.[3]

At the end of 1960 and beginning of 1970, Professor Hershko became interested in how proteins are degraded in cells because there is an intrinsic energy-dependence of intracellular protein degradation.[7] To start revising how degradation of

proteins takes place, Hershko and Ciechanover studied the mode of action of fraction of reticulocyte lysates. Two main fractions were classified, the fraction 1 related with those proteins not adsorbed to a DEAE-cellulose resin and the fraction 2 compiled those proteins adsorbed to the DEAE-cellulose resin. After different purification approaches including heat treatment, a small protein ATP-dependent was the key to elucidate the mode of action of the proteolytic system. This small protein, first called AFP-1 later Ubiquitin, was studied to see its binding for instance, to a protein present in the Fraction 2 of the reticulocyte lysate. According to the results, the binding was by a covalent bond, specifically an amide linkage, since there were no high-molecular-weight derivatives when alkali, hydroxylamine or SDS (boiling) conditions were present when the protein was incubated even in the presence of β-mercaptoethanol.[8]

Ubiquitin (Ub), an essential protein in eukaryotes, is a 76 amino acid protein highly conserved from yeast to plants and mammals. It is globular and compact. It has one α-helix (from residues 23-34), a 3_{10} helix (from residues 56-59), and five β-strands: β1 (residues 1-7), β2 (residues 11-17), β3 (residues 40-45), β4 (residues 48-50) and β5 (residues 65-72).[8] Function-relevant features of Ubiquitin's compact 3D structure include i) a hydrophobic patch surface along residues L8, I44 and V70; ii) some electrostatic potential due to the positive environment created around the hydrophobic patch by residues K6, R42, K48, H68 and R72; iii) a flexible C terminus; iv) seven lysine residues (K6, K11, K27, K29, K33, K48, K63) and the N terminus which could be used for polymerization (Figure 2).[9]
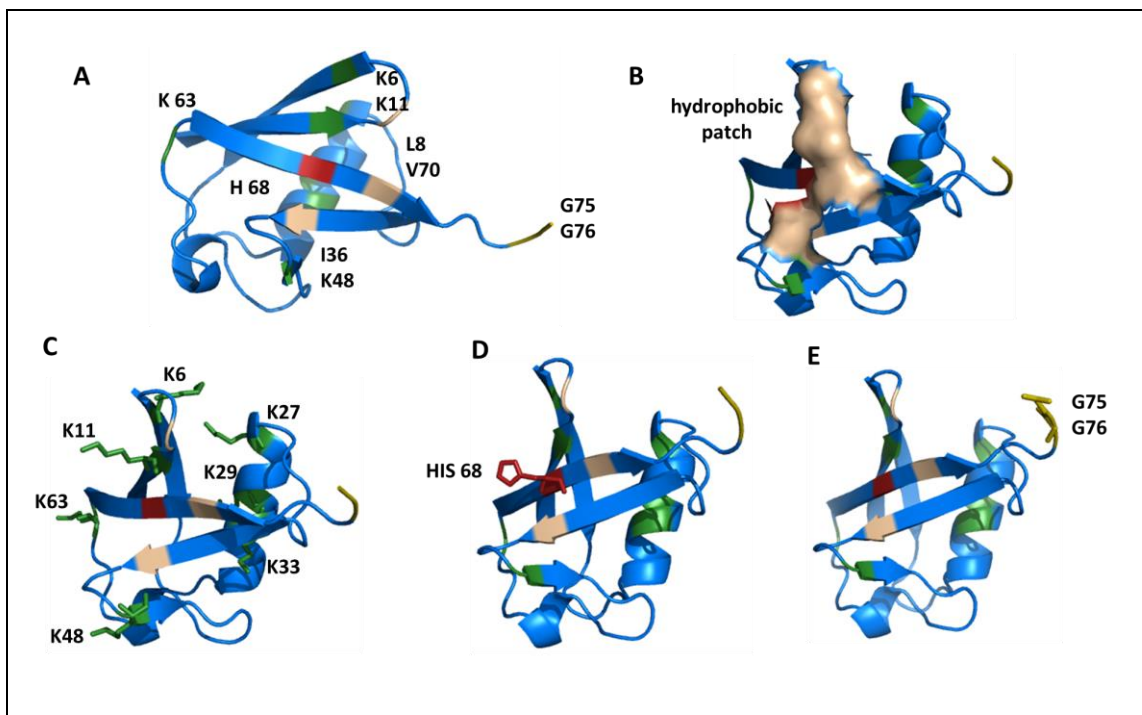
Figure 2. Ub 3D ribbon structure depiction; A: ubiquitin structure model, B: hydrophobic patch surface, C: lysine residues, D: His68, pH sensor, E: glycines 75, 76 C-terminal motif.

Further studies lead to the hypothesis that Ubiquitin (Ub) is a protein tag. At the beginning it was thought that Ub is attached by some enzymatic mechanism to a target protein, and then reused after other enzymatic hydrolysis process. This hypothesis was proved right and it is summarized in Figure 3.[10]
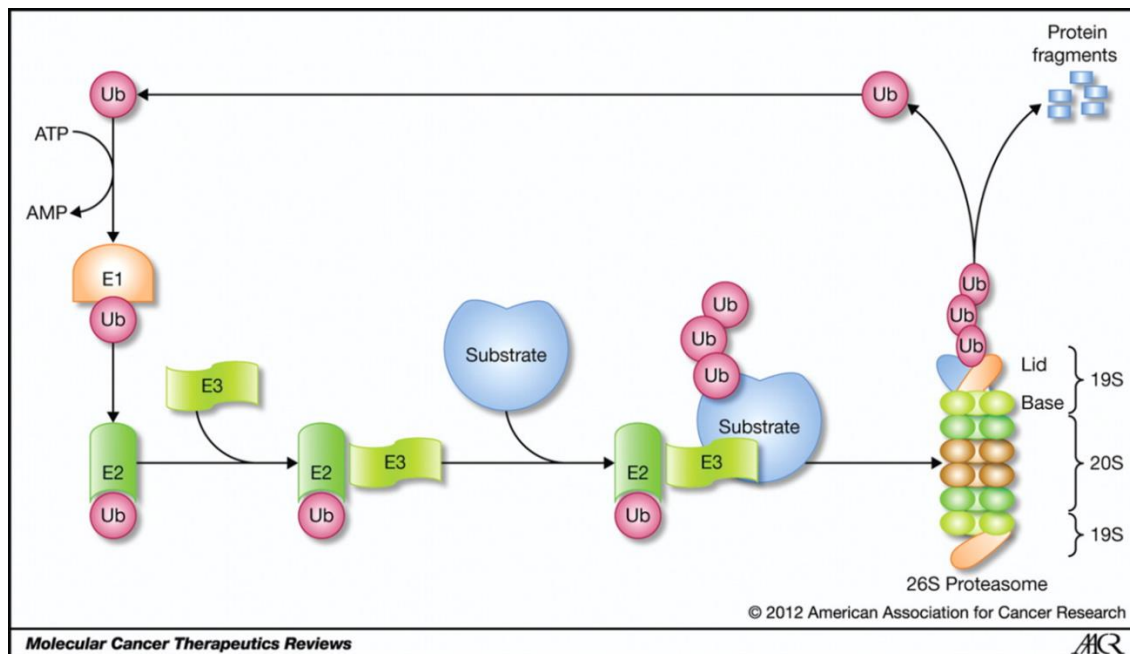
Figure 3. Ub proteasome system pathway.[10]

The covalent ligation of Ub to a protein occurs through an isopeptide bond. Chemically, the isopeptide bond is an amide bond. The amide functional group is ubiquitous since it is neutral, stable and it has hydrogen-bond accepting and donating properties. The formation of an amide bond is a condensation reaction between an acid and an amine which has to overcome a thermodynamic barrier as the equilibrium lies on the side of an acid-base reaction rather than a synthesis of the amide bond, as shown in Scheme 1. A direct condensation could be achieved at high temperature (160–180 ˚C), and as a consequence it is quite incompatible. Therefore, the activation of the carboxylic acid as a leaving group is mandatory.[11]
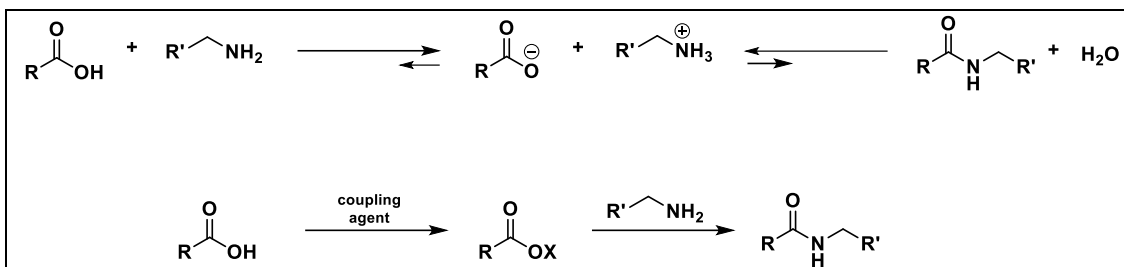
Scheme 1. Amide bond formation. Top: the unfavorable thermodynamic equilibrium; bottom: the strategy to overcome the energetic barrier and favor the synthesis of the amide bond.

The amide bond is achieved by the lysine residue in protein synthesis. Lysine has in its structure two primary amines, and to differentiate which amine group is used to achieve the amide bond the convention dictates the α-amino group will form a peptide bond whereas the ε-amino group will form an isopeptide bond, as depicted in Scheme 2.
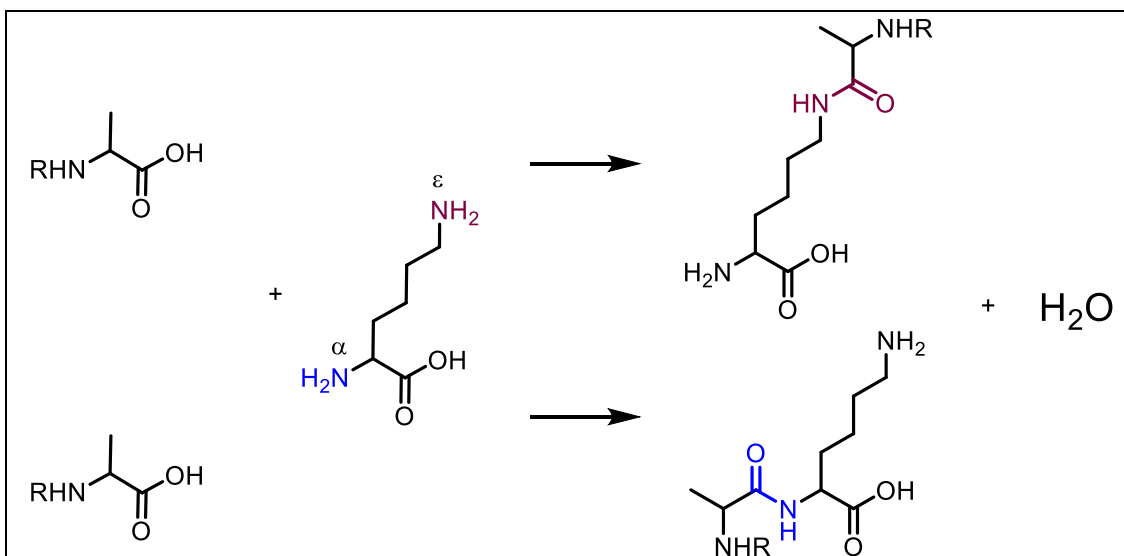
Scheme 2. Peptide - Isopeptide bond. By the α-amino group the amide is called peptide bond (blue) whereas by the ε-amine it is referred as isopeptide bond (red).

In nature, the thermodynamic barrier for protein synthesis is reduced since the activation process is mediated by catalytic enzymes. Ubiquitination is not an exception. It requires the sequential action of three enzymes called the E1-E3 system.

Enzyme E1 activates the C-terminal glycine residue with ATP. An immediate intermediate Ub-adenylate is formed. After the release of a pyrophosphate, Ub becomes bonded through a thioester to a free cysteine residue in E1. By a transthioesterification reaction, Ub is then transferred to the E2 Ub-conjutating enzyme. In the third step of reaction, the E3 enzyme (Ub ligase) links through an amide bond the C terminus of Ub to the side chain of a lysine residue of another protein. This amide bond received the name of isopeptide linkage (Scheme 3).[12]
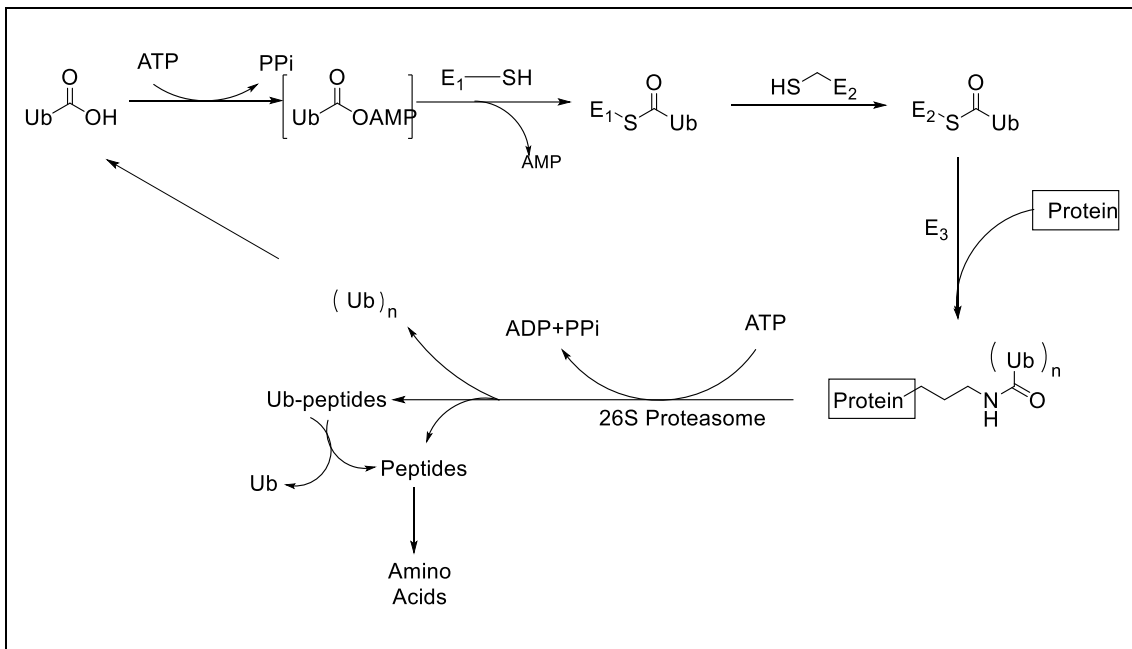
Scheme 3. Ub activation pathway.

When a protein is tagged by Ub, the process is called ubiquitination. Critical signaling pathways are controlled by this tagging, and in eukaryotes this control is essential for cell proliferation, differentiation and survival.[13, 14] The target protein then might be affected by the attachment of Ub in terms of its cellular localization, activity, interaction partners and structure.[15] Basically two modes of ubiquitination exist as summarized in Table 1.[16, 17]

Type of Ubiquitination	Features/Characteristics
Monoubiquitination	Attachment of one Ub to a single lysine residue in the target protein. The modified target protein usually shows different interactions, localization or activity. Multiple monoubiquitination occurs when several lysine residues of a target protein have a single Ub molecule attached.
Polyubiquitination	Attachment of a Ub chain to a lysine residue in the target protein. A Ub chain exhibits a hydrophobic patch motif which is required for binding to receptors to further drive the target protein toward its final fate. In addition, when a chain is present in the target protein, two forms of polyubiquitination are known: homogeneously linked Ub and mixed-linkage Ub chains which could be of linear topology or branched. The biological function signal depends on the lysine residue of Ub used to build the Ub polymer.

Table 1. Types of ubiquitination

Monoubiquitination is a regulatory process of the location and activity of diverse cellular proteins. It is involved in at least different cellular functions, such as histone regulation, and endocytosis.[17] The best studied examples of polyubiquitination (homogeneous) are when the chain is built by Ub moieties linked through the K48 residue or K63 residue. First, K48 polyubiquitin chains are the signal that targets a substrate protein toward the proteasome for degradation.[10] Studies conducted by Nuclear Magnetic Resonance (NMR) and X-ray crystallography revealed that K48-linked chains have a preferred closed conformation in which the hydrophobic surface patch residues of two adjacent Ub molecules form an interface.[9] By contrast, the polyUb chains built through the K63 residue do not act as a signal for degradation but, in parallel with monoubiquitination, generate a non-proteolytic signal involved in DNA repair, transcriptional regulation, endocytosis and activation of protein kinases.[18] In terms of the conformational appearance K63 chains adopt an extended arrangement in solution in which the hydrophobic patches are not in closed contact.[19]

Moreover, Ub signal is perceived by specific domains (UBDs) which are carried by proteins that represent signal transducers. The latter proteins are called Ub receptors and they constitute a Ub-based network of protein-protein interactions.[16] A comprehensive classification of these domains can be summarized in four classes (Table 2).[20]

Domain class	Known members	Features
Helical domains	UBA (ubiquitin associated) UIM (ubiquitin-interacting motif) DUIM (double-sided UIM) MIU (motif interacting with Ub) CUE (coupling of Ub conjugations to endoplasmic reticulum degradation) GAT [GGA (Golgi-localized, gamma-ear-containing) and TOM (target of Myb)]	All are known to interact with a single region on Ub, the Ile44 hydrophobic patch.
ZnF (Zn Fingers)	NZF A20 ZnF UBP-ZnF	Discovered through dissections of known Ub-binding proteins. There are three regions on the surface of Ub which are recognized. The affinities range from 1μM to mM.
Ubc (Ubiquitin-conjugation enzyme)-related domains	UEV (Ubc E2 variant)	Known as intermediates between the activating enzyme and the ligase, E1 and E3, respectively.
Other UBDs	UBM (ubiquitin-binding motif) Jab1/MPN domain	Binding is centered at the Ile44 hydrophobic patch. Binding can take place in mono-ubiquitin and in polyubiquitin.

Table 2. UBD classification summary.

What happens with Ub after the target protein has been processed by its tagging, in a simplified way, is that it gets recycled and reused in the cell; since as mentioned earlier, ubiquitination is a tight dynamic and reversible regulated enzymatic process. For this particular purpose, detaching Ub or hydrolyzing a polyUb chain is done in nature by another group of enzymes called deubiquitinating enzymes (DUBs). DUBs are first peptidases; in other words they catalyze the cleavage of the bond formed between Ub-Substrate, and Ub-Ub. Second, DUBs are also classified in five different families based on their homology of their catalytic domains (Table 3). Third, four of these families are cysteine proteases and just the JAMM family is a group of metalloproteases. And as known, the catalytic mechanism of the cysteine protease DUB works in a triad; in other words, a histidine residue will arm the cysteine for nucleophilic attack on the carboxylic center of the peptide bond while the aspartate aligns and polarizes the histidine. The metalloproteases, in contrast, require the coordination of a Zn ion for catalysis in such a way it will allow the abstraction of a hydrogen from water making a hydroxyl ion the nucleophile which will attack the electrophilic center of the peptide bond.[21]

Family	Known members	Features
Ubiquitin C-terminal hydrolase (UCH)	UCHL1, UCHL3, UCHL5, BAP1	Cysteine proteases
Ubiquitin-specific protease (USP)	USP1, USP2, USP3, USP4, USP5, USP6, USP7, USP8, USP9X, USP9Y, USP10, USP11, USP12, USP13, USP14, USP15, USP16, USP18, USP19, USP20, USP21, USP22, USP24, USP25, USP26, USP27X, USP28, USP29, USP30, USP31, USP32, USP33, USP34, USP35, USP36, USP37, USP38, USP39, USP40, USP41, USP42, USP43, USP44, USP45, USP46, USP47, USP48, USP49, USP50, USP51, USP52, USP53, USP54, DUB3, CYLD, USPL1	
Ovarian tumor (OTU)	OTUB1, OTUB2, OTUD1, OTUD3, OTUD4, OTUD5, OTUD6A, OTUD6N, OUT, YOD1, Otulin, A20, Cezanne, Cezanne2, TRABID, ACPIP1	
Machado-Joseph disease (MJD)	Ataxin-3, Ataxin-3-like, JosD1, JosD2	
JAB1/MPN/Mov34 (JAMM)	PMSD7, PMSD14, EIF3H, BRCC36, C5CN5, AMSH, AMSH-LP, MPND, PRPF8, MYSM1	Metalloproteases

Table 3. Families of deubiquitinating enzymes and their members

Ultimately, in the ubiquitination process the most important feature is the formation of an isopeptide bond. This enzymatic reaction cascade is a constant challenge due to the vast collection of enzymes in order to achieve a site specific linkage.[13] Therefore, it is important to expand a toolbox of chemical methods for site-specific installation of Ub and Ub-like modifiers and their application to study how ubiquitination takes place.

In the present dissertation, the research was focused on the application of suitable organic chemistry reactions from a possible construction of a semisynthetic ubiquitin to isopeptide bond mimics. In order to shine some light in the structural facts when Ub is part of a polymeric chain in these conditions NMR spectroscopy was used as the characterization technique. Hence, this dissertation is divided in four chapters. The first chapter describes the chemical synthesis of lysine analogs and their use to be included in the Ub polypeptide chain. The second portrays the efforts to characterize Ub dimers built by a silver catalyzed condensation where the nucleophilic side chains are orthogonally protected. In the third chapter, the construction of Ub dimers with a non-canonical isopeptide bonds is described and examined if the pH has a direct effect on their hydrophobic interfaces. Finally, in the fourth chapter, future work beyond the present scope is explored.

Chapter 1: Synthesis of Mercaptolysine Analogs as Unnatural Amino Acids in Ubiquitin Pursuing Native Chemical Ligation Auxiliaries

Part A. Preliminary and representative examples of lysine surrogates

As mentioned before, building Ub chains or ubiquitination of a substrate protein depend on the E1-E3 enzyme system, which can be a difficult task since not all enzymes are known. And the golden prize is to achieve the isopeptidic bond in the most natural way possible. Activation of the carboxylic acid consists of the replacement of the hydroxyl group with a suitable leaving group since the acid would otherwise just have an acid-base reaction with the amine. Therefore, the coupling reaction to form an amide bond requires an efficient mechanism that is appropriate for the substrates based on their stereochemistries and conformations. Most of the precedents, to bypass the E1-E3 enzyme system, with a semisynthetic approach to study ubiquitination (as later exemplified) are related with building a chain of Ub where the isopeptide bond has a non-hydrolyzable functional group to bond covalently with a DUB or a UBD in order to elucidate how affinity/binding takes place. The attempts for building a chain of Ub by synthetic methods are a combination of synthetic peptides built by solid phase peptide synthesis (SPPS) and ligation methodologies focused to have quantitative and homogeneous starting materials for characterization. The selected examples describe "protein-friendly" chemical methodologies in terms of not using harsh organic compounds which could

compromise the stability of the protein or not full recovery of the folded protein structure. They also served as a starting point to address a way for a non-enzymatic attempt to attach Ub either to another Ub or to a protein substrate; as a consequence the description of each methodology is directed toward the chemical approach used.

I.   Thioether ligation

Ub was built by the conjugation of a peptide bearing a cysteine residue as a nucleophilic system and a N-chloroacetyl peptide as an electrophile to produce by alkylation a thioether which mimics a lysine residue.[22] Later, expanding the application, the methodology was used to build a K63-linked Ub dimer, as shown in Scheme 4.
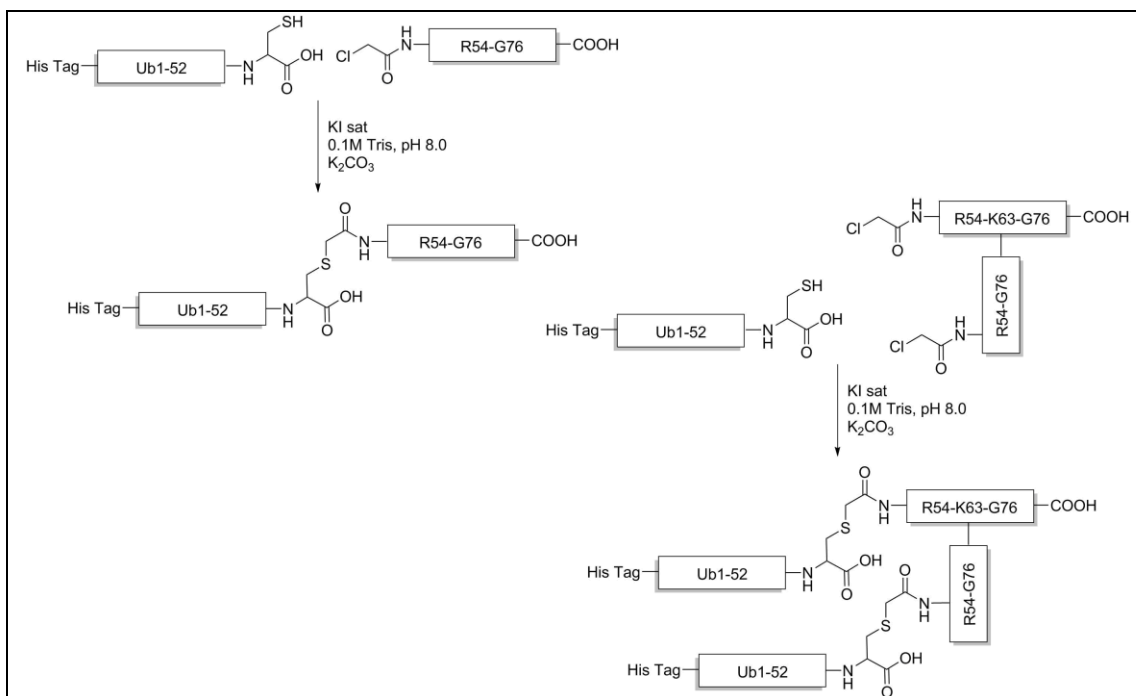
Scheme 4. Ligation reaction between a cysteine and N-chloroacetyl to produce a thioether in substitution of a lysine residue.

II.   Oxime ligation

By SPPS and a reverse trypsinolysis,[23] Ub was ligated to small peptides from PCNA, H2B, H2A and FANCD2 which are known to be ubiquitinated. The lysine residue on these peptides is a synthetic aminoxyacetyl-L-diamino propionic acid; therefore, a non-hydrolyzable Ub is achieved and used for DUB affinity as shown in Scheme 5.
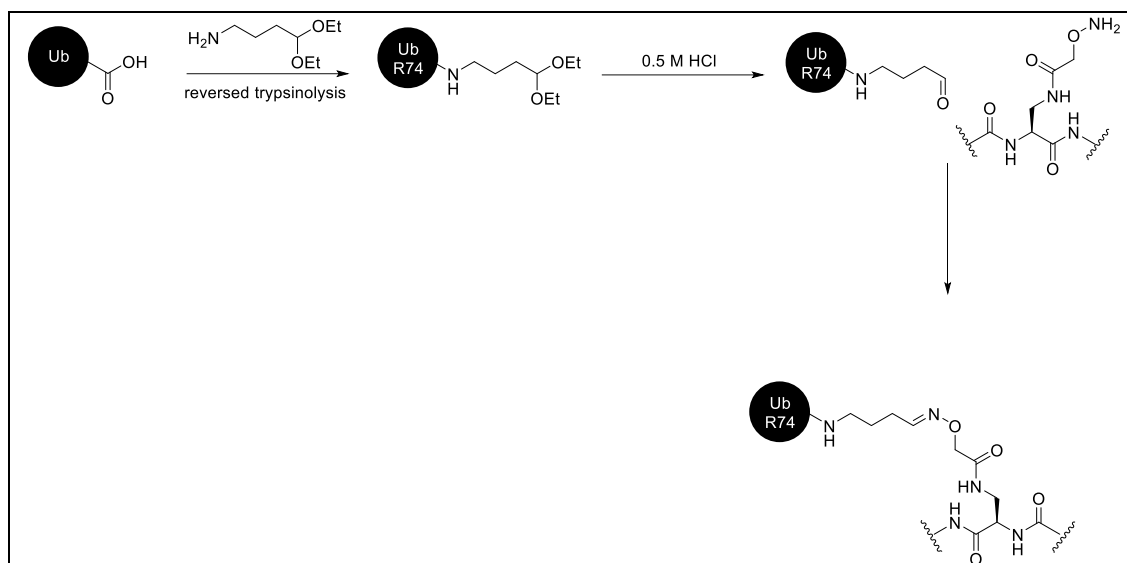
Scheme 5. Oxime ligation

III.   Pseudoproline and Dimethoxybenzyl building blocks.

An intricate linear Fmoc-SPPS with pseudoproline building blocks as well as dimethoxybenzyl dipeptides Ub was built in order to prevent aggregation intermediates in the purification with the type of resin they used.[24] Six positions in Ub were identified for optimal incorporation of the building blocks (Scheme 6). After detachment from the resin, further purification and characterization by circular dichroism (CD) yielded a protein with the same profile as a recombinant protein and it reacted enzymatically with the same specificity toward the formation of polyUb chains.
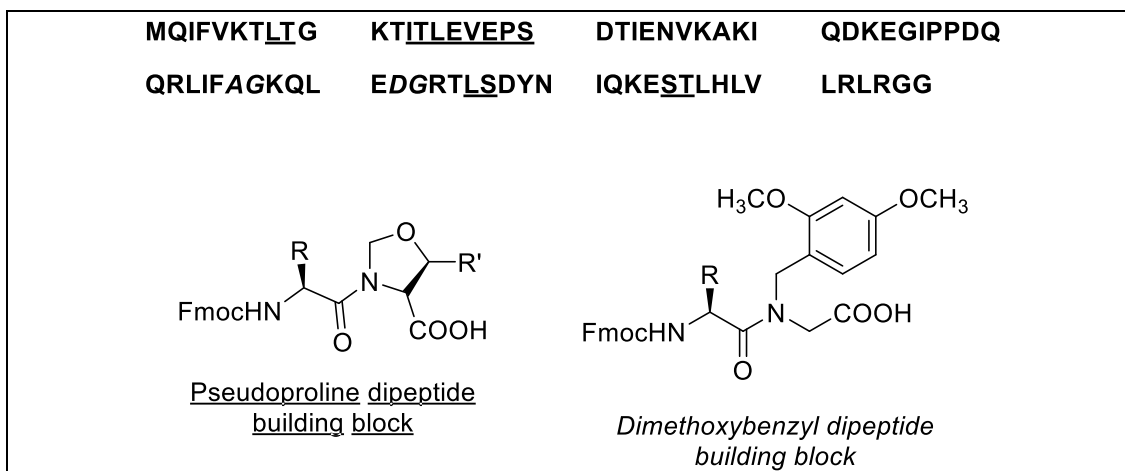
Scheme 6. Synthesis of Ub by Fmoc-SPPS with pseudoproline and dimethoxybenzyl building blocks.

IV.   C terminus derivatization by Intein Chemistry

The description of a methodology about how to modify the C terminus of Ub and other Ubls was used to target the specificity of DUBs.[25] Two Michael-addition substrates, one acts as a reversible inhibitor (vinyl sulfone) and the other as an irreversible one (aldehyde) were used to react by an addition-elimiation reaction with Ub thioester (UbCOSR) which was isolated when 2-mercaptoethane sulfonate (MesNa) was used to cleave Ub from an intein-chitin bead domain (CBD) affinity tag, as shown in Scheme 7.
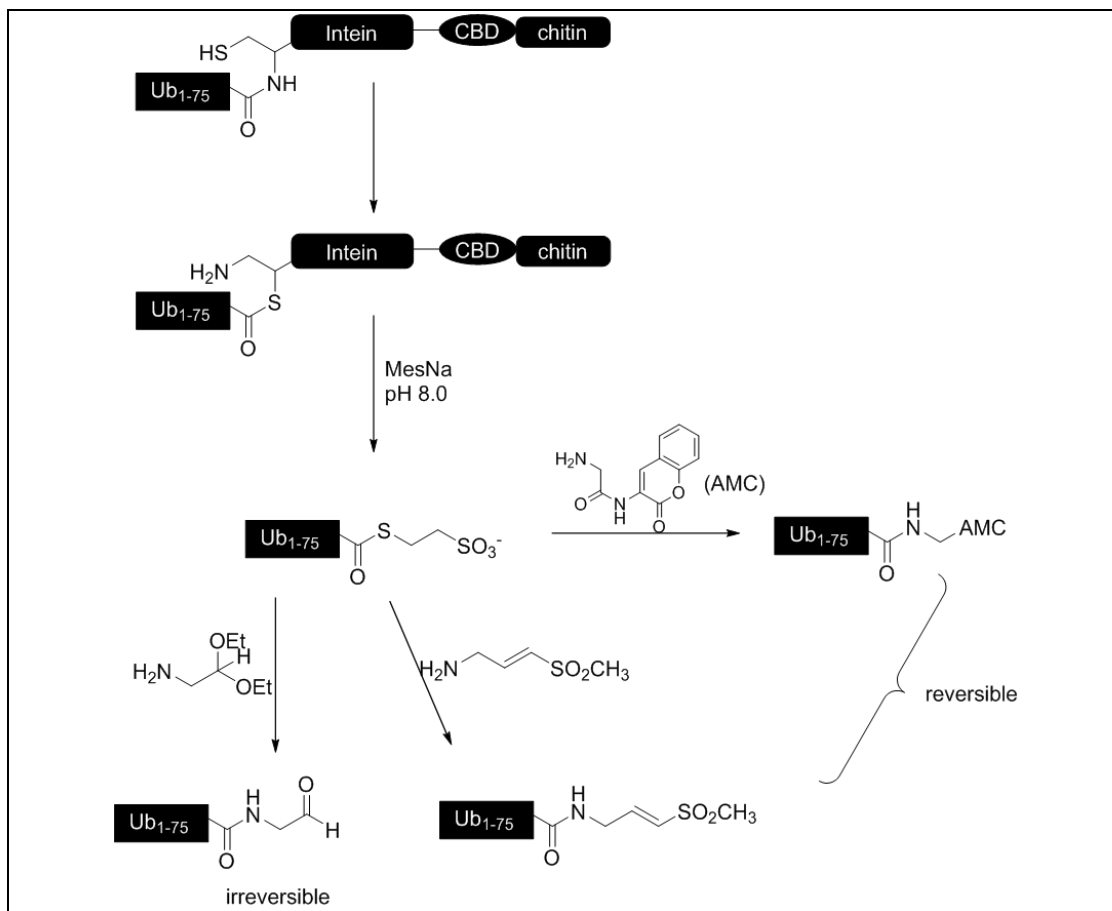
Scheme 7. Addition-elimination reaction with reversible/irreversible inhibitors to UbCOSR isolated from Intein-CBD affinity tag.

V.   Previous mercaptolysine syntheses

a.   Synthesis of δ-mercaptolysine

In 2009, Brik's research group first reported the synthesis of δ-mercaptolysine starting from glutamic acid.[26] After fifteen steps this lysine analog was incorporated into the Ub structure by SPPS (Scheme 8). Then, it was used to ubiquitinate a small peptide containing the fragment of α-synuclein. The methodology then evolved towards a better protection of the ε-amino group (Scheme 9).[27] Later, Chin's research group reported the synthesis of the same residue starting from δ-hydroxylysine for its incorporation into Ub by site-specific unnatural amino acid mutagenesis methodology (Scheme 10).[28]
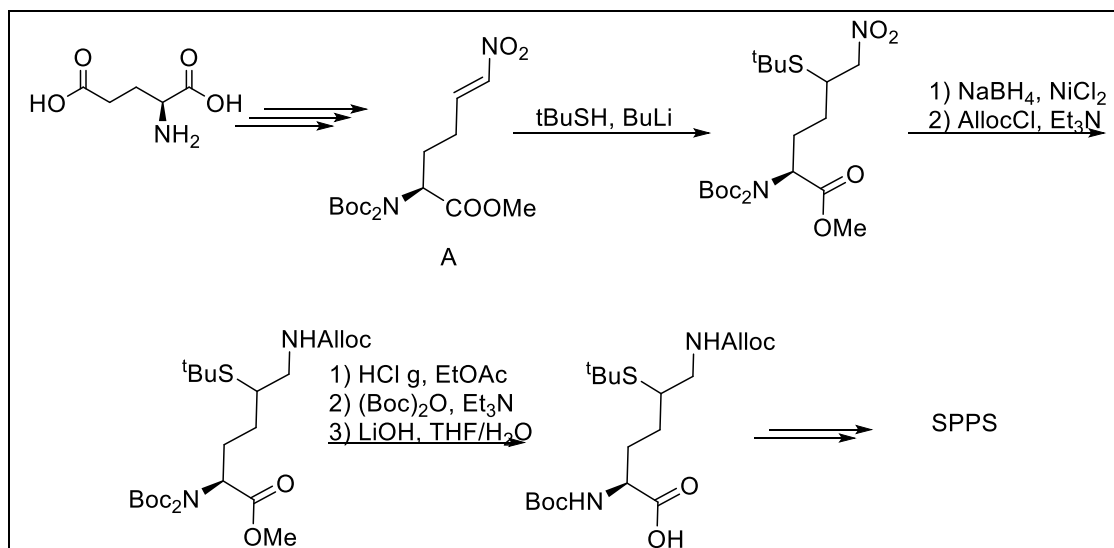
Scheme 8. Synthesis of δ-mercaptolysine
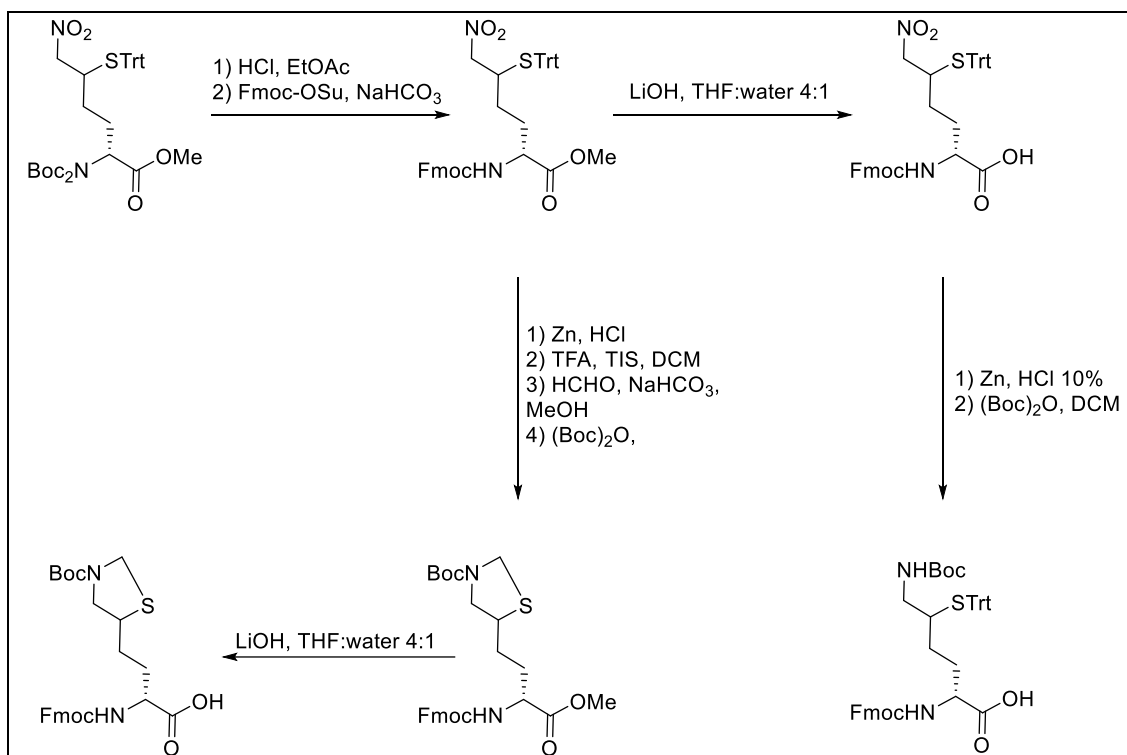
Scheme 9. Alternative ε-NH₂ protecting groups in δ-mercaptolysine
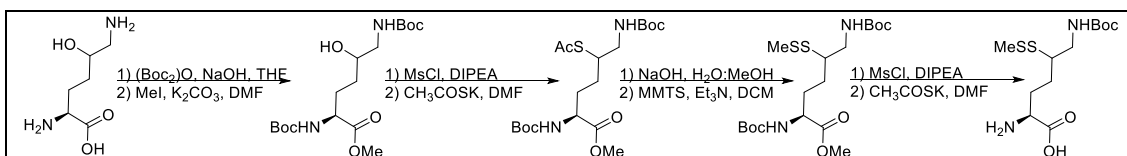
Scheme 10. Synthesis of δ-mercaptolysine for site specific unnatural mutagenesis

b. Synthesis of γ-mercaptolysine

Also in 2009, Liu's group developed the synthesis for the γ-mercaptolysine to be used for SPPS;[29] starting from aspartic acid. In a linear synthesis of eighteen steps, including the ring opening of a 4,6-dioxopiperidine precursor,[30] a thiol handle is located in a symmetric manner because it is in the middle of the lysine backbone. To incorporate the mercaptolysine in the sequence of Ub, first an orthogonal deprotection of the ε-NH₂ was performed. Then, by native chemical ligation (NCL) with a

fragment of Ub took place. Finally, the amino acid functionality was deprotected and

in particular, a second NCL took place with biotin (Scheme 11).
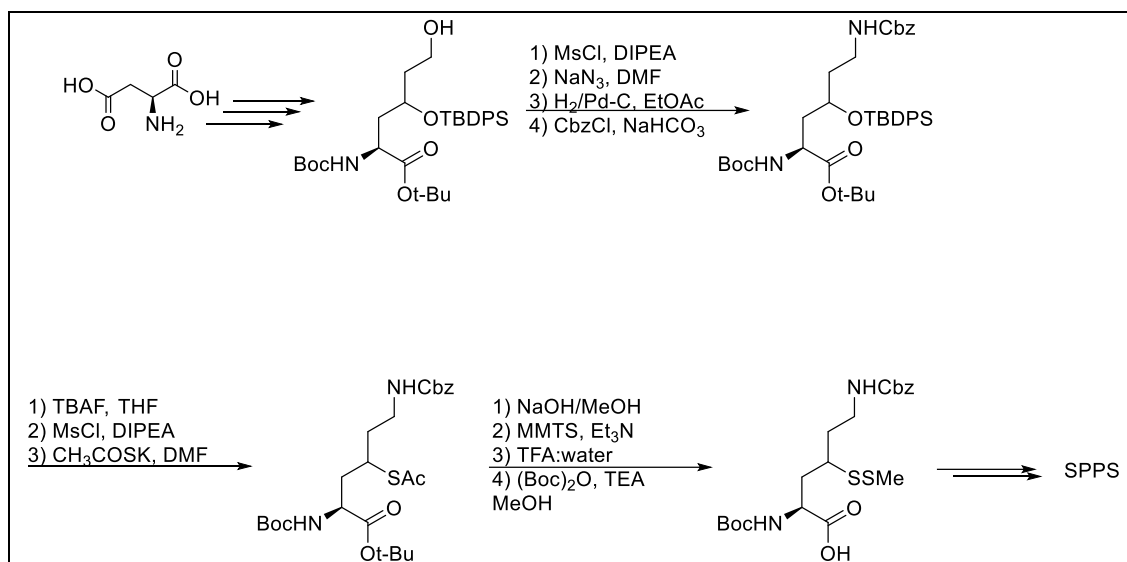
Scheme 11. Synthesis of γ-mercaptolysine

Part B. Alternative synthesis of mercaptolysines as native chemical ligation

auxiliaries

From the previous examples one can extract the information that there is a number of chemical strategies which can be applied without compromising the integrity of the protein, since the second most important target to achieve is that the protein can fold into the right tertiary structure and that the isopeptide bond is not rigid.

The starting point of the mercaptolysine project was to take advantage of some features the structure of such unnatural amino acid residue could offer as a target for achieving ubiquitination in an enzyme-free fashion. First of all, the possibility of improving the synthesis of such amino acid was an attractive starting point. Second, NCL in combination with site-directed mutagenesis was the driving force since this amino acid could have been incorporated at any position in Ub's polypeptide chain. And finally after ligating Ub with a target protein, it allowed the possibility to characterize it by NMR.

But what is NCL? Concisely, the definition of NCL is the reaction between an activated carboxylate (thioester) and a 1,2-aminothiol which allows the formation of native backbone structures chemoselectively (Scheme 12A). The first step of the reaction is the nucleophilic addition of the sulfhydryl group to the activated carboxylate by a transthioesterification.

Then, the RDS (rate determining step) occurs when the primary amine makes a secondary addition reaction in an intramolecular and spontaneous rearrangement, or more precisely called, an S→N acyl transfer occurs. As a consequence an amide bond is achieved and the product of reaction bears a free sulfhydryl group which can be later removed, if necessary.[31] Now, in peptides and proteins, the thioester is best accomplished either by direct attachment of such leaving group or by cleaving the protein from the expressed intein tag, generally attached to the protein by recombinant methodologies, with the thioester of preference such as MesNa or MPAA. And the 1,2-aminothiol generally is a free cysteine residue.[32]

Likewise, site-directed mutagenesis of unnatural amino acids is the methodology of choice when more chemical diversity needs to be added to a protein. The methodology involves the use of three types of plasmids. One encodes for the tRNA which recognizes a stop codon, the second encodes for the evolved amino acyl RNA synthestase (aaRS) and the third one encodes the protein of interest with the stop codon mutation (most commonly used: the amber TAG codon) at a specific position which will lead to the incorporation of an unnatural amino acid (Scheme 12B). The first and second plasmids are necessary to avoid the recognition of endogenous tRNAs and block the protein formation when cells are transfected with such plasmids; this concept is called orthogonality and the best way to satisfy this criterion is when the non-endogenous tRNA synthetase comes from a distant species of the host tRNA synthetases. The host machinery expresses the protein of interest with the unnatural amino acid placed instead of any of the other 20 canonical amino acids because the medium is enriched with it.[33]
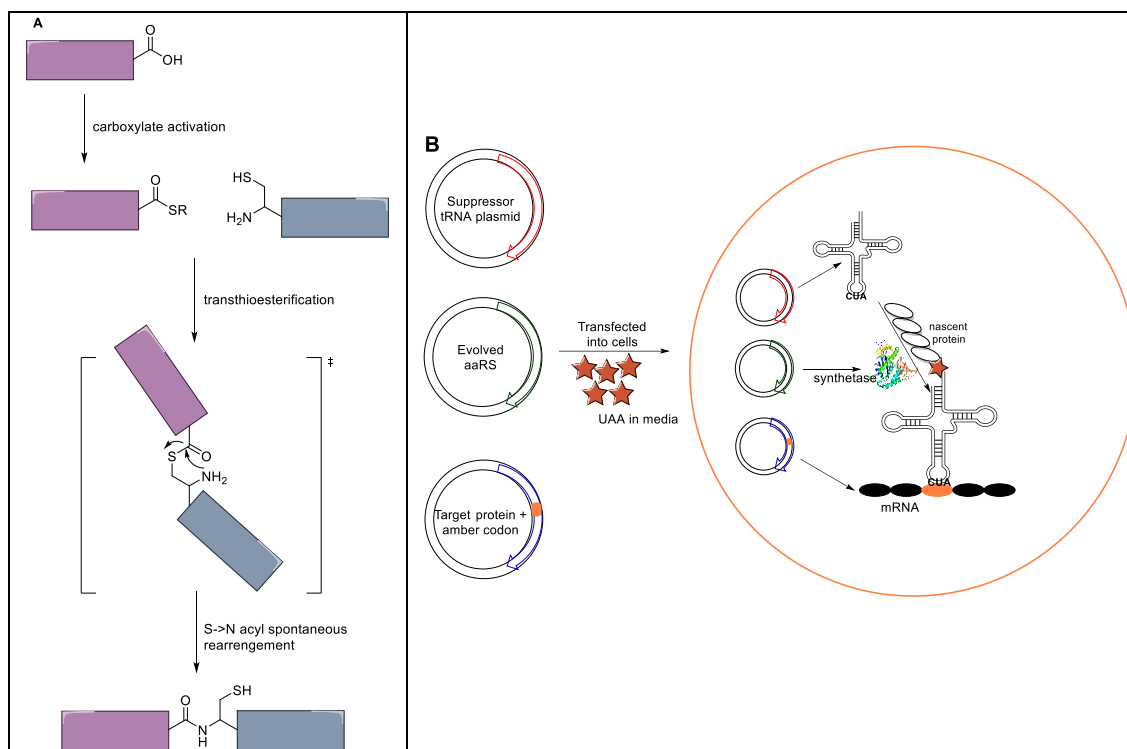
Scheme 12. A: Native Chemical Ligation strategy[31]; B: Unnatural amino acid site-directed mutagenesis[33]

Initially, the main objective was to reproduce the synthesis of γ-mercaptolysine done by Liu *et al* for its incorporation in the polypeptide chain of Ub as an unnatural amino acid. However, this linear synthesis gave a low yield of reaction after the ring opening of the dioxopiperidine precursor. Thus, the synthesis of this amino acid was a starting point to develop an improved and in house procedure.

a. Synthesis of γ-mercaptolysine, (γSH)K

In the retrosynthetic analysis of the lysine backbone of γ-mercaptolysine **6** there are 4 crucial steps for developing the synthesis (Scheme 13). **A** corresponds to the S_N^2 reaction of sodium azide to a carbon center bearing a good leaving group. **B** is the most important intermediate because it is an α,β-unsaturated aldehyde which calls for a 1,4-Michael addition donor system. **C** is an intermediate which can lead to an easy way to elongate the side chain toward the lysine backbone. And finally **D** is aspartic acid, one of the canonical 20 amino acids and a very cheap starting material.
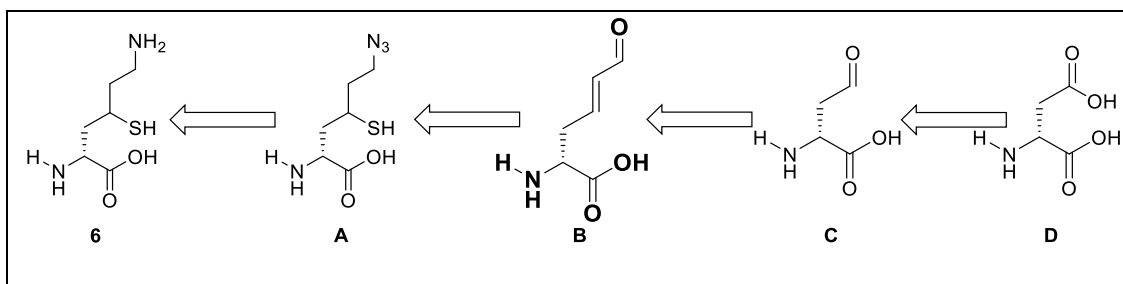
Scheme 13. Retrosynthetic analysis for γ-mercaptolysine.

As Brik used a γ-aldehyde in his synthesis of δ-mercaptolysine from glutamic acid;[27] the γ-aldehyde from aspartic acid was also achieved previously, according to Padron *et al* report.[34] Therefore, going from **D** to **C** is a regioselective reduction, but only if the reaction conditions are done exclusively in anhydrous toluene conditions. A Wittig-Horner reaction is the best procedure to go from **C** to **B**. Previously Sutherland reported the synthesis of an analog of **B**, but it was highlighted that the Wittig-Horner only proceeds if the α-amine has one *tert*-butyl carbamate.[35]

Thus, following these reaction conditions, **1** was synthesized and was the starting point towards the improved synthesis of the γ-mercaptolysine (Scheme 14). Finally to wrap up the idea of the retrosynthetic analysis, compounds **3** and **4** are the analogs of **A**.
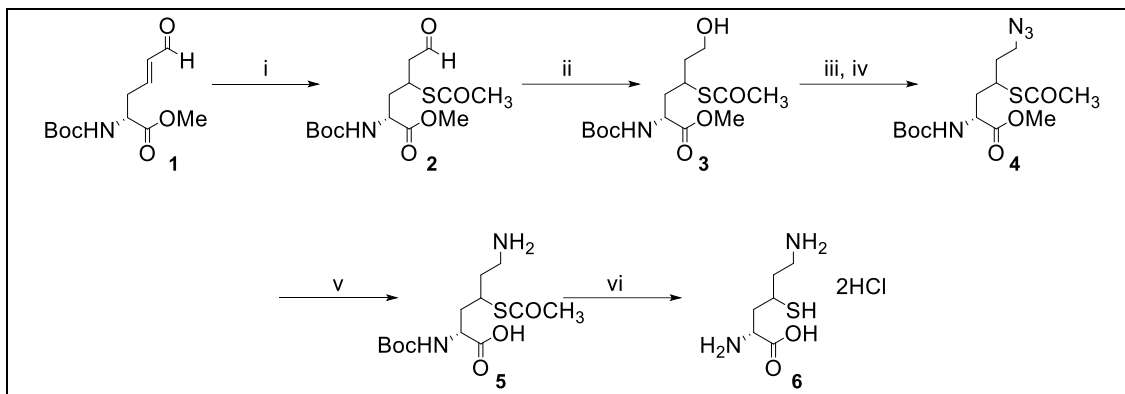
Scheme 14. Synthesis of γ-mercaptolysine. i) CH$_3$COSH, H$_2$O:MeOH (40:60), V-50 (10% mol), 50 °C; 75%. ii) THF, NaBH$_4$, -10 °C; 80%. iii) DCM, DIEA, MsCl, 0 °C; 92%. iv) DMF anh, NaN$_3$, 80 °C; 91%. v) THF anh, PPh$_3$, 60 °C; 85%. vi) MeOH:H$_2$O (2:1), NaOH 1N; HCl (12N, dioxane), 0 °C; 68% (two steps).

However, the 1,4-Michael addition of thioacetate to **1** was not an easy task because the yield of reaction dropped after trying different experimental conditions where the conventional addition-elimination mechanism was attempted. At this stage the reaction from **1** to **2** was crucial and the most important step in the synthesis. Therefore, all the effort was directed to developing the best reaction conditions towards the thiol-Michael addition using potassium thioacetate or thioacetic acid as the -SH handle precursor.

First, the 1,4-Michael addition of the thioacetate moiety to aldehyde **1** by using a buffered system of potassium thioacetate and thioacetic acid (AcSH) in three different solvent conditions (Table 4) was attempted.[36] Due to the solubility of **1**, polar aprotic solvents such as DMF and DMSO left a heterogeneous solution at 10 °C. Aprotic polar solvents such as ethyl ether and THF gave better results. However, the best conditions were obtained when the aldehyde **1** was directly dissolved in AcSH, followed by the addition of potassium thioacetate (Table 4, entry 3). On average, **2** was obtained after purification with a 70% yield.

Solvent	Yield (%)
Et$_2$O	50
THF	30
AcSH	70
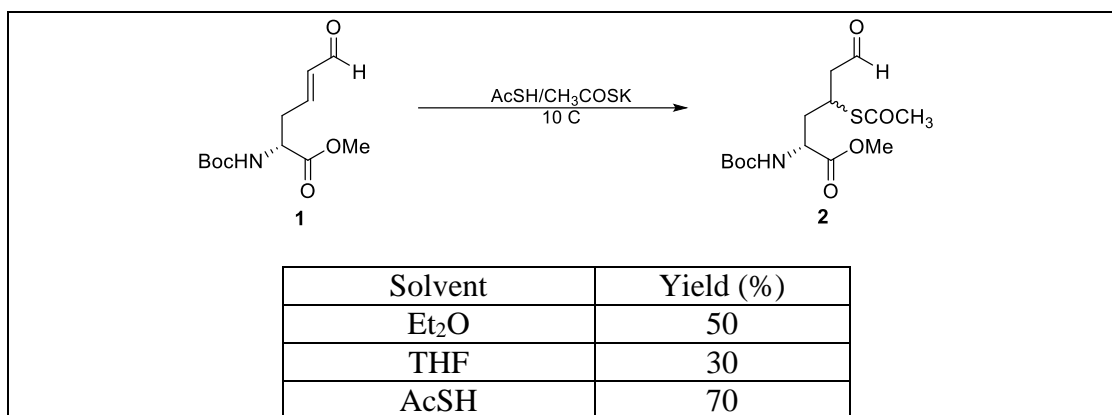
Table 4. Solvent screening for conjugate addition

Aldehyde **1** in comparison with other α,β-unsaturated systems, reviewed for thiol Michael addition,[37] lacks of an electron-withdrawing group making it not as electrophilic and as a consequence less reactive.[38] Other experimental conditions were considered, such as the addition of AcSH via a free-radical process.[39] Water-soluble azo-initiators have been exploited in polymerization radical reactions under mild conditions with traces of side products.[40]

Thus, the addition of AcSH to **1** in the presence of the two water-soluble radical initiators, VA-044 and V-50 was carried out. Considering that the reaction involves the abstraction of hydrogen from some source, the solvent system selected was where aldehyde **1** remained in homogeneous solution. Therefore, a mixture of two protic solvents such as water-methanol was used. The thermal azo-initiators were used first in a stoichiometric fashion at their respective half-life decomposition temperature (Table 5). The highest yield of **2** was obtained in a 60:40 methanol:water ratio with V-50 as initiator (72% yield as a 1:1 diastereomeric mixture) after purification. With these conditions selected, the amount of V-50 initiator needed for the reaction was optimized. The screening conditions were from 1% mol to 100% mol (stoichiometric) of V-50. The yield of reaction increased to 75% when the molar ratio of initiator to aldehyde **1** was 10% mol (Table 6). It is very important to emphasize that no special conditions such as inert atmosphere were needed and that the reaction can be done in aqueous solution.

Table 5. Methanol-Water screening for free radical conjugate addition

Azo-initiator	Temperature (C)	CH$_3$OH:H$_2$O	Yield (%)
VA-044	44	50:50	60
VA-044	44	60:40	70
VA-044	44	70:30	50
V-50	50	50:50	40
V-50	50	60:40	72
V-50	50	75:25	60

C$_{V-50}$ (%mol)	Yield
1.0	no reaction
5.0	45
10	75
50	70
100	72

Table 6. Azo-initiator free radical concentration screening

To conclude the synthesis of **6,** the reduction of the aldehyde **2** with NaBH$_4$ in wet THF to avoid the reduction of the thioester moiety was carried out.[41] Alcohol **3** was substituted with an amino group, first by replacing the alcohol group with an azide to give **4**, which was then reduced by the Staudinger reduction to give the amine **5**.[42] A shortcut in the synthetic strategy could be generated from **2** by reductive

amination,[43] however, this procedure was not pursued due to the need of ammonium gas which is highly toxic. After alkaline and then acidic hydrolysis, γ-mercaptolysine **6** was isolated with an overall yield of 28% after six linear steps beginning from **1**. With this approach the number of steps, relative to the previously reported synthesis,[44] was reduced and the free radical Michael addition was the most important optimization step in the synthesis sequence. Moreover, the free radical reaction was used in a different application than the polymerization processes.

b. *Synthesis of δ-mercaptolysine*, (δSH)K

Following the same approach of a Michael acceptor as a common starting point, the synthesis of δ-mercaptolysine was revised. In the same context as in γ-mercaptolysine the following retrosynthetic analysis which satisfies the synthetic route can be proposed (Scheme 15).
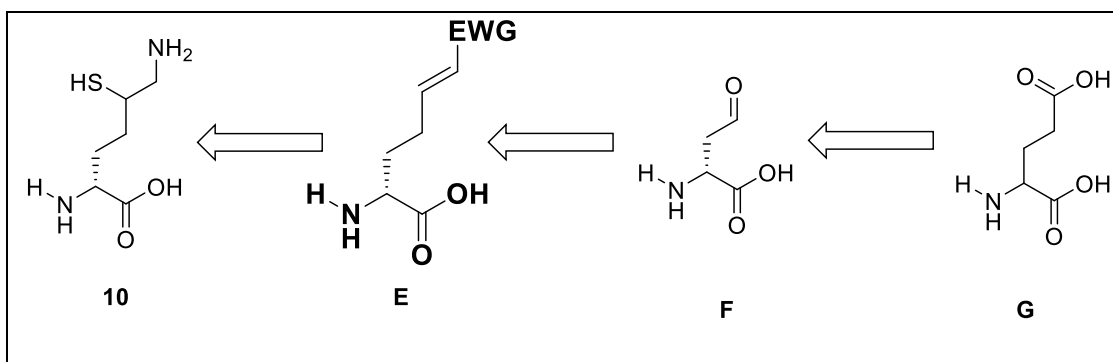
Scheme 15. Retrosynthetic analysis for δ-mercaptolysine

The synthesis of δ-mercaptolysine **10**, began with the racemic nitro olefin **7**, which is the analog of **E** because it bears the nitro as an electro withdrawing group (EWG). The Michael addition to nitro olefin **7**, previously reported,[27] was

accomplished by the direct addition of potassium thioacetate, a straightforward thiol precursor, taking advantage that it is hydrolyzed using mild alkaline conditions. The use of potassium thioacetate has been reported for the Henry reaction to be successful with high yields.[45] The conjugate addition of potassium acetate to the nitro olefin **7** was successfully achieved in high yield (90%) and the attachment of the thioester gave a racemic mixture of **8** (Scheme 16).

To avoid the reduction of the *tert*-butoxycarbonyl (Boc) group to an oxime, acidic deprotection of both Boc groups was performed first with HCl in dioxane followed by the addition of Zn to afford the corresponding amine **9**, in a one-pot reaction. Finally, the total deprotection of the thiol, as well as the carboxylic acid moiety, took place by hydrolysis in alkaline and then acidic conditions to give the δ-mercaptolysine **10**. As a result, the synthesis of **10** was reduced by two synthetic steps that previously reported by Brik *et al*, [27] and the overall yield was 42-45% over the four steps starting from **7**.
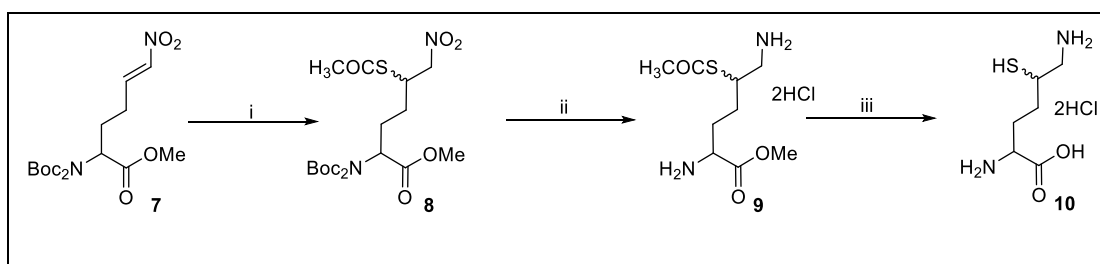
Scheme 16. Synthesis of δ-mercaptolysine; i) Et$_2$O, CH$_3$COSK, -15 °C; 90%. ii) HCl (12N, dioxane), 25 °C; Zn powder 10% mol, 25 °C; 70%. iii) MeOH:H$_2$O (2:1), NaOH 1N.

Once the amino acids were synthesized, their incorporation into the sequence of Ub was attempted; and for that purpose site-directed mutagenesis was the methodology selected.

Part C. Site-directed Mutagenesis

The first attempt for protein expression with the mercaptolysines was using GFP as the reporter system. The rationale for this change was that GFP fluoresces when UV light is irradiated. Therefore, if the incorporation of any mercaptolysine is successful, the detection of the expression is straightforward. As explained earlier, an exogenous (orthogonal) tRNA synthetase is needed for the incorporation of an unnatural amino acid in the desired position of the protein sequence, and the systems listed in Table 7 were used. Here, it is very important to mention that in order for the orthogonal pyrrolysine/tRNA synthetase to incorporate any of mercaptolysines, they must bear a carbamate in the ε-NH$_2$. This pair incorporates into protein's sequence unnatural amino acids related to pyrrolysine in response to the amber stop codon TAG, and this system is useful when the recombinant protein is expressed, for instance in *E. coli*.[46] Thus, the ε-NHBoc derivatives of γ- and δ-mercaptolysine (**12** and **13**, respectively, Figure 4) were selected,[47] since the positive control responds with N-ε-Boc-Lysine.
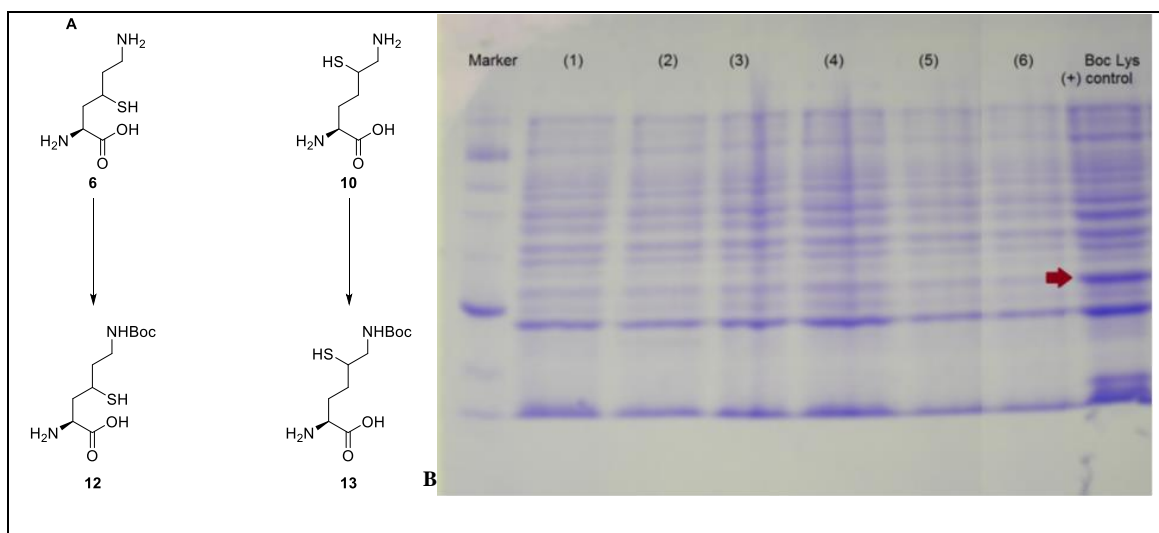
Figure 4. GFP protein expression with mercaptolysines; A: ε-NH2 carbamate analogs of γ- and δ-mercaptolysine. B: SDS-PAGE protein expression test for the pPylsY384F/sfGFP pair; Lanes (1): No UAA (2): 6 (3): 12 (4): 10 (5): 13 (6): Lysine

aaRS/reporter	Lysine analog	Expression Fluorescence	aaRS/reporter	Lysine analog	Expression Fluorescence
pPylsY384F/ sfGFP	Boc-Lys	Y	pSupONB N436A/ pTrcHisA Y151A sfGFP	Boc-Lys	N
	δ-SHK	N		δ-SHK	N
	γ-SHK	N		γ-SHK	N
pBKev17/ sfGFP	Boc-Lys	N	pSupONBLys /pTrcHisA Y151A sfGFP	Boc-Lys	N
	δ-SHK	N		δ-SHK	N
	γ-SHK	N		γ-SHK	N
pSupInterm+/ pTrcHisA Y151A sfGFP	Boc-Lys	N	pSupInterm /pTrcHisA Y151A sfGFP	Boc-Lys	N
	δ-SHK	N		δ-SHK	N
	γ-SHK	N		γ-SHK	N

Table 7. Protein expression tests with exogenous tRNAs synthetases and GFP as reporter protein

From the unsuccessful protein expression tests as listed in Table 7, one possible way to solve the lack of incorporation of the mercaptolysine amino acids could be to create mutations in the synthetases until the amino acid could be recognized. However, the δ-mercaptolysine incorporation into Ub was published by Chin's research group, using a variant of the pPylsY384F from *M. bakeri*.[28]

Nevertheless, this opened another approach to be implemented. It is known that 1,2-aminothiols are useful for NCL because the S→N acyl transfer is somewhat a spontaneous process.[31, 48] So far in the literature the methodology is considered only for synthesis of proteins by SPPS or semi-recombinant techniques.[31, 49] But, what could happen if these mercaptolysines instead of being considered amino acids that mandatory have to be part of the polypeptide chain instead of the canonical lysine are an addition to the polypeptide sequence that naturally would not be part of the protein? Moreover, is it possible to attach them directly to Ub's structure and then characterize any interesting outcome?

Part D. Native Chemical Ligation with Ub

The attachment of mercaptolysines **6** and **10** <u>in solution</u> at the C terminus of Ub by NCL was tested with the purpose of enabling the introduction of an intrinsic spin-label system or a fluorophore, further detailed in Part E, without introducing a mutation in the amino acid sequence of the protein. This way, the C terminus of Ub would be attached through an isopeptide bond to the ε-NH$_2$ of the lysine or, as a second possibility, through a regular peptidic bond if the NCL occurs at the α-NH$_2$ of the amino acid functionality, as previously depicted in Scheme 2. The presence of the spin-label at the C-terminal tail of Ub would allow mapping long-distance interactions by means of NMR.

Strategically, first Ub's G76 COOH was functionalized as a thioester by the catalytic reaction of the E1 enzyme in the presence of ATP, MgCl$_2$ and MesNa (Scheme 17)[24, 50] in order to activate the carboxylic acid with a suitable leaving group.

After lyophilization, UbCOSR **11** was solubilized in the NCL buffer at pH 7.2 and incubated with each mercaptolysine (**6** and **10**; Scheme 18A) at 25 °C in non-denaturing conditions for the protein. After cation exchange chromatography (Scheme 18D) the ligation product yield was 80%, and the correct mass was confirmed by ESI-MS (Scheme 18B).

Scheme 17. Thioesterification of Ub C terminus. A: reaction scheme, B: ESI-MS deconvoluted spectra of the protein; top: Ub expected 8564 Da, found 8562.6 Da, bottom: UbCOSR expected: 8687 Da, found 8688.7 Da.
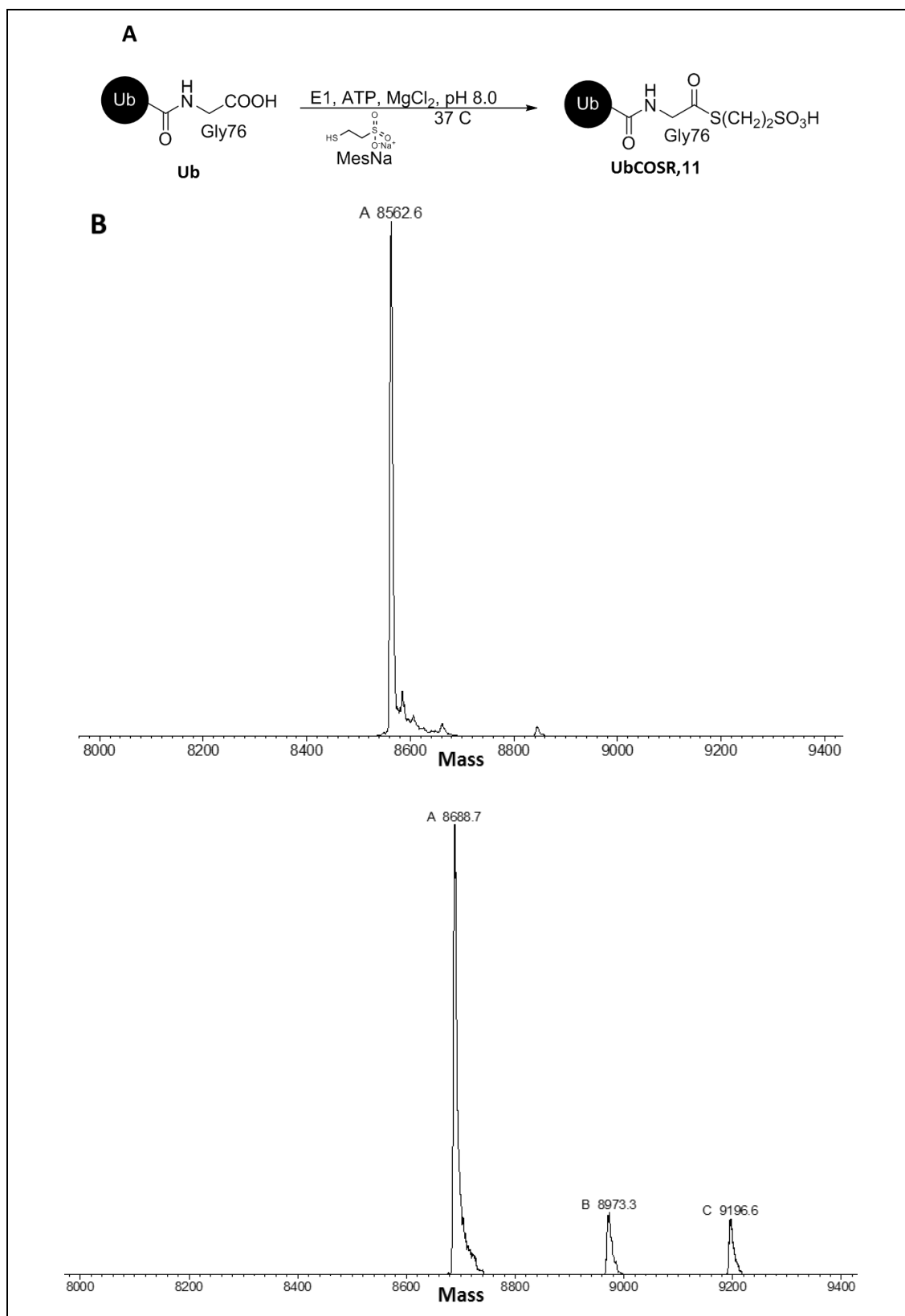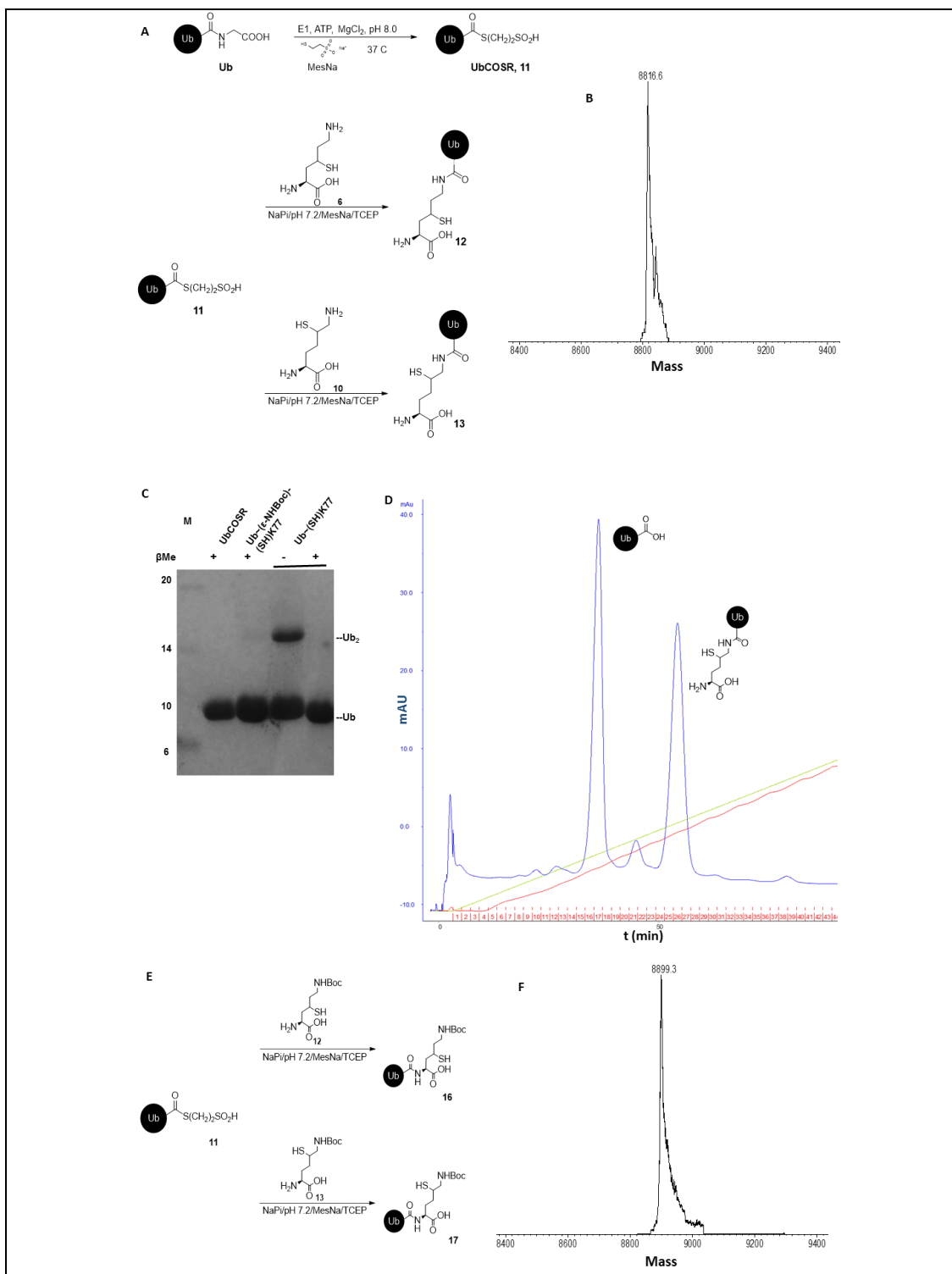
Scheme 18. Native Chemical Ligation of Mercaptolysines with Ub. A: Synthesis of Ub~(SH)K77; B: ESI-MS deconvoluted spectrum of Ub~(SH)K77; C: SDS-PAGE of the NCL reaction; D: High-performance cation exchange chromatography of the NCL reaction; E: Synthesis of Ub~(ε-NHBoc)-(SH)K77; F: ESI-MS deconvoluted spectrum of Ub~(ε-NHBoc)-(SH)K77. A and E reaction conditions: 200mM Na$_2$HPO$_4$ pH 7.2, 100mM MesNa, 60mM TCEP, 25 °C, 80% yield.

The SDS-PAGE analysis (Scheme 18C) revealed that in the absence of β-mercaptoethanol as disulfide reducing reagent some of the reaction products migrate as a Ub dimer. However, under reducing conditions the dimer disappeared, indicating that the Ubs were connected via a disulfide bond and not by an amide bond. In addition, the reactions with the ε-NHBoc analogs (**12** and **13**) showed the addition of one Ub unit (Scheme 18E and 18F). As expected, the NCL occurs at the level of the α-amino group since it is the most reactive toward the S→N acyl shift.[48] Considering the Baldwin's rules,[51] and the ring formation favored by the Thorpe-Ingold effect, because the array is the most substituted,[32] the regioselectivity in the cyclic transition state is established by a 6-membered ring for the γ-mercaptolysine for "both ends" (Scheme 19) and a 5-membered ring (ε-NH$_2$) preferentially over a 7-membered ring (α-NH$_2$) for the δ-mercaptolysine (Scheme 20).
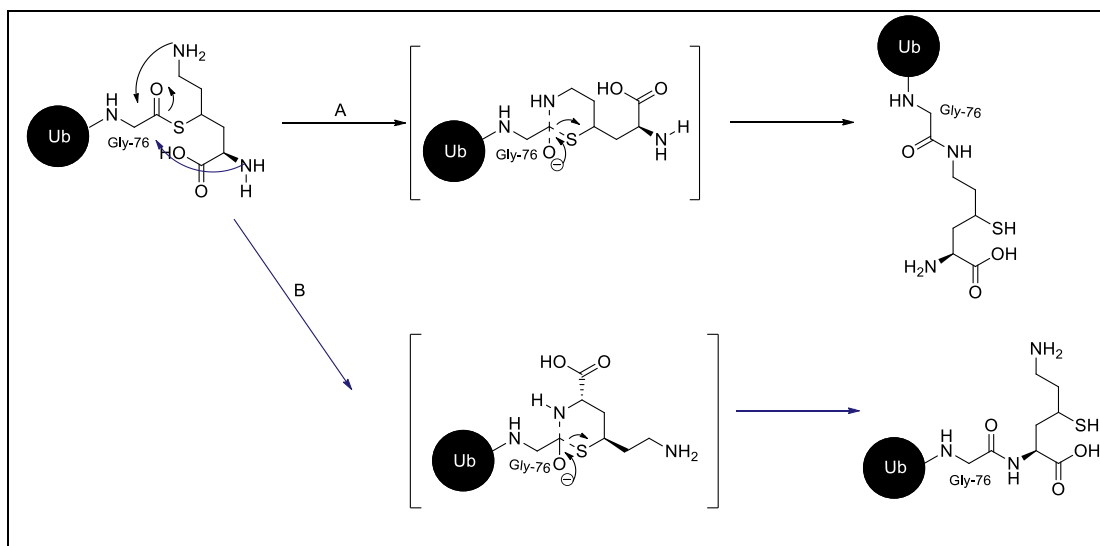
Scheme 19. (γSH)K NCL acyl shift curved-arrow mechanism
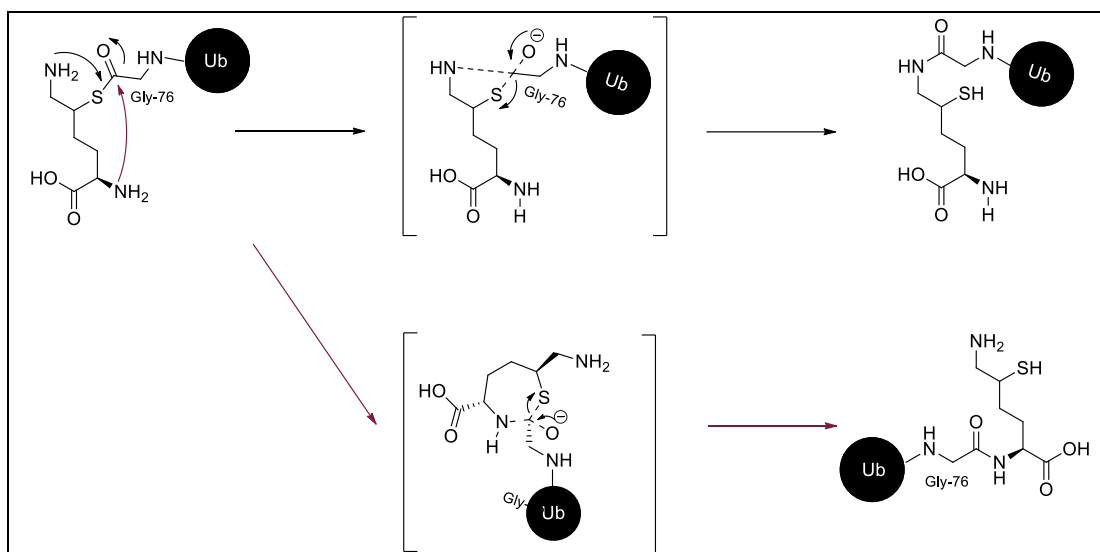
Scheme 20. (δSH)K NCL acyl shift curved-arrow mechanism

To corroborate that the addition of the mercaptolysine is completely reversible a simple enzymatic hydrolysis reaction of Ub~(SH)K77 (**12** and **13**) with a Ub C-terminal hydrolase (YUH1) at slightly alkaline conditions (pH 8.0) for a period of four hours was performed.[52] After analysis by ESI-MS, Ub was recovered with no trace of the previously attached mercaptolysine (Scheme 21).

With this partial characterization, Ub~(δSH)K77 (**12**) was used for the fluorophore and paramagnetic effect studies. A more thorough characterization by NMR was done toward the alpha vs epsilon linkage preference due to the fact that the NCL was done as previously stated in solution and with both amino groups exposed for S→N acyl shift.
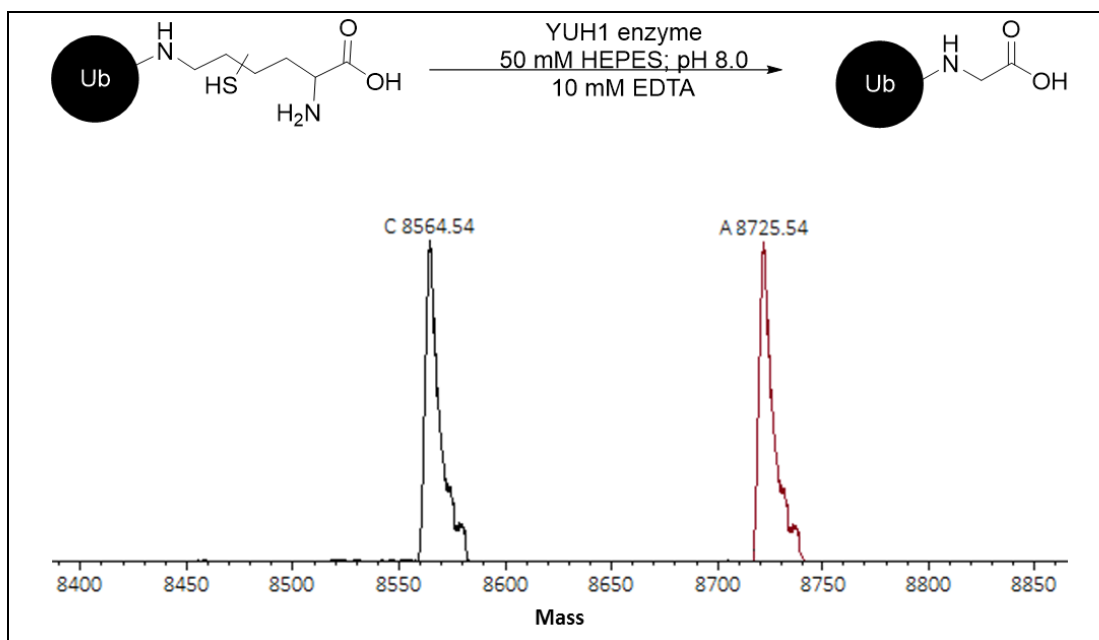
Scheme 21.YUH1 Hydrolysis reaction of Ub~(SH)K77; A: (SH)K77Ub 8725.54 Da, C: Ub 8564.54 Da

Part E. Fluorophore attachment

As mentioned before, the attachment of the mercaptolysine residue to Ub by NCL gives access to further use the intrinsic thiol handle, since it is a nucleophilic system. Thus, first as a test of availability, fluorescein-5-maleimide was attached (Figure 5) in particular with Ub~(δSH)K77, **12**, at pH 8.0.[53] The use of fluorescent tags is useful when complex mixtures of proteins need to be resolved. According to the deconvoluted ESI-MS spectrum since no **12** was detected, the yield of reaction is higher than 90%. Therefore, the attachment of (SH)K allows the use of fluorescence labels to trace Ub in various mixtures.
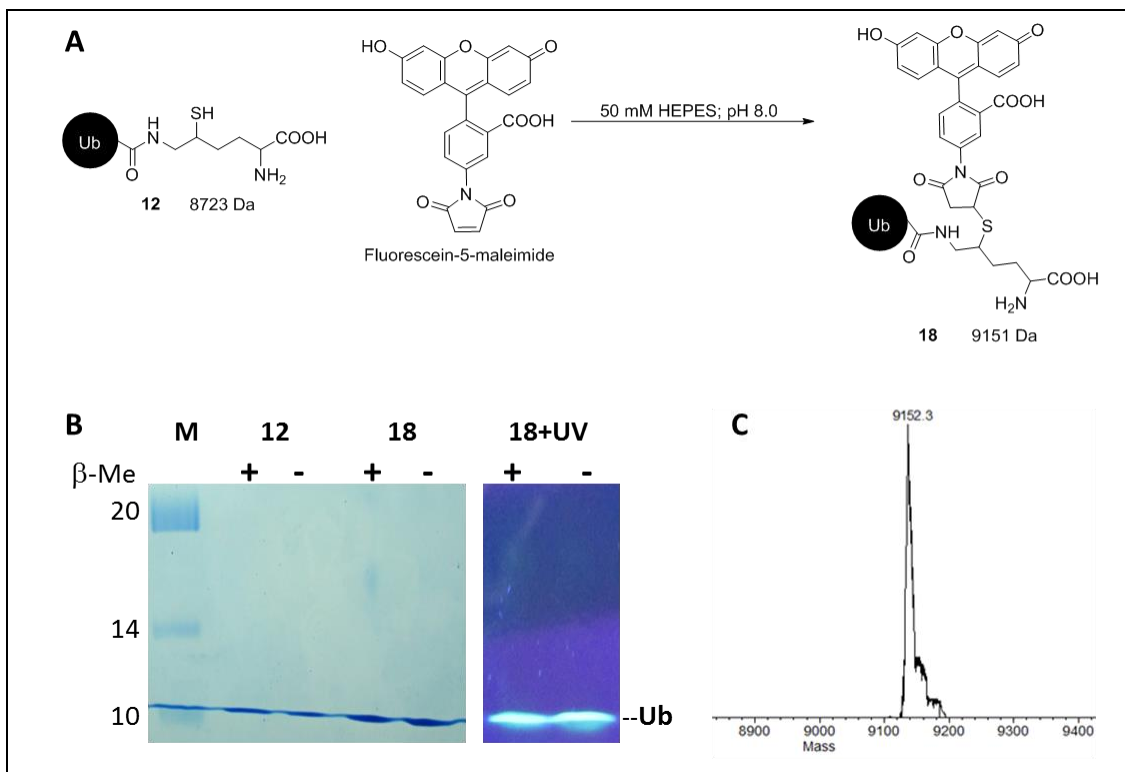
Figure 5. Fluorescein-5-maleimide attachment to Ub~(δSH)K77. A: Reaction strategy. B: SDS-PAGE of the fluorescent labeling, Coomasie staining at left and under UV light; the samples were incubated with 8M urea to avoid false dimer bands. C: Deconvoluted ESI-MS.

Part F. NMR characterization and paramagnetic relaxation enhancement effect

To determine if the presence of the mercaptolysine residue has an effect on Ub a series of ^{15}N-^{1}H HSQC NMR spectra were collected; using wild type Ub (^{15}N-Ub^{WT}) as reference (Figure 6) for any overlay and any calculation based on its chemical shifts of each amide signal.
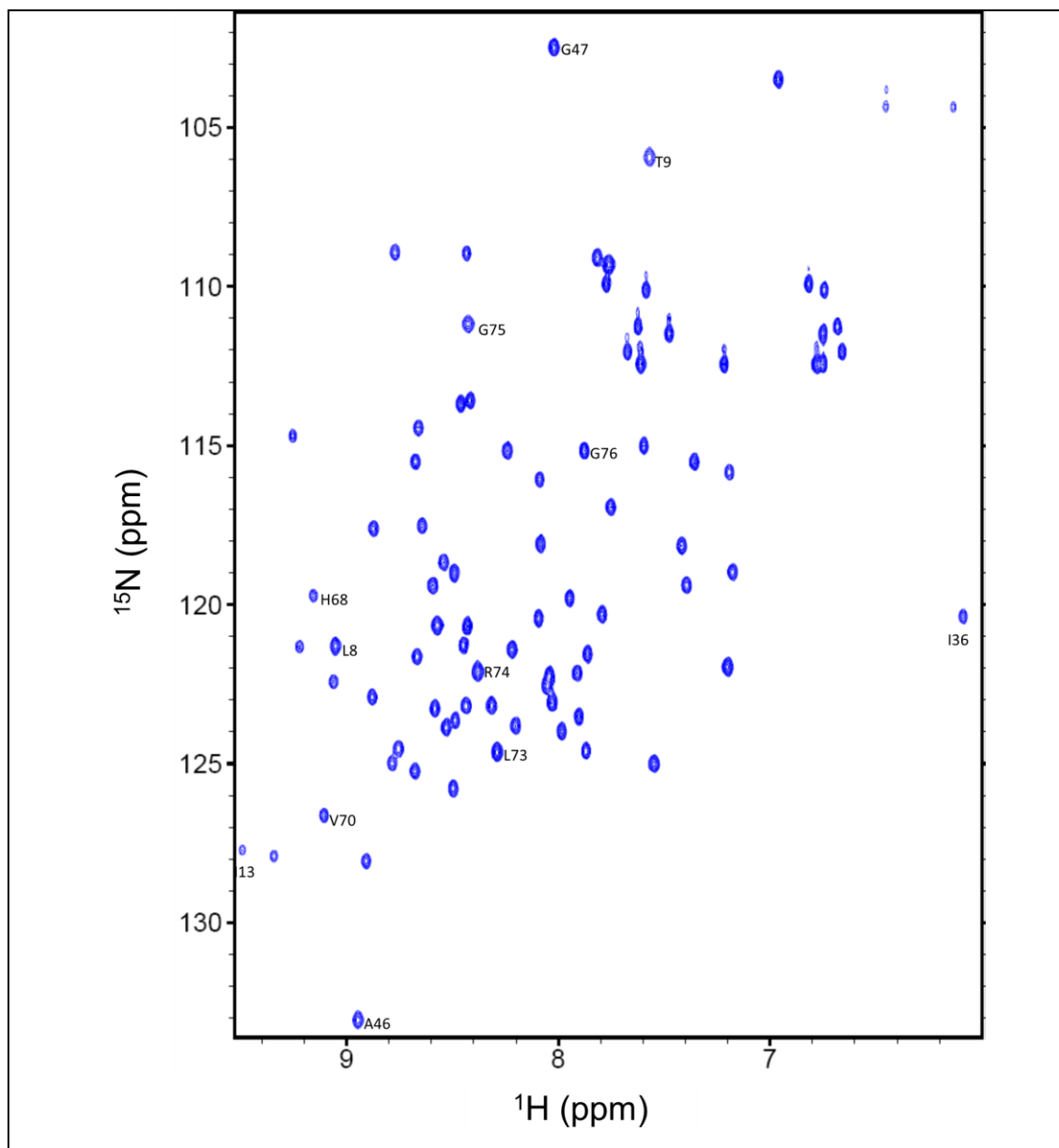
Figure 6. ^{1}H-^{15}N HSQC NMR spectrum of ^{15}N-Ub^{WT}. Some amide signals are labeled by residue letter and number following Ub polypeptide sequence

To begin, the spectrum of ^{15}N-Ub(δSH)K77 was recorded. Since the amino acid is attached at the C terminus of Ub, the rationale suggests that the addition of the mercaptolysine amino acid will have the highest effect at the tail. The effect is quantifiable in terms of a perturbation of its chemical shift (Figure 7B). In other words, the amide signal of the neighboring residues would move in the ppm-scale. And effectively, the mercaptolysine residue has a larger effect from residues L73 to G76 (Figure 7A), with the latter exhibiting the strongest effect. This is fully consistent and directly indicates that (δSH)K77Ub is covalently attached to the C terminus of Ub.
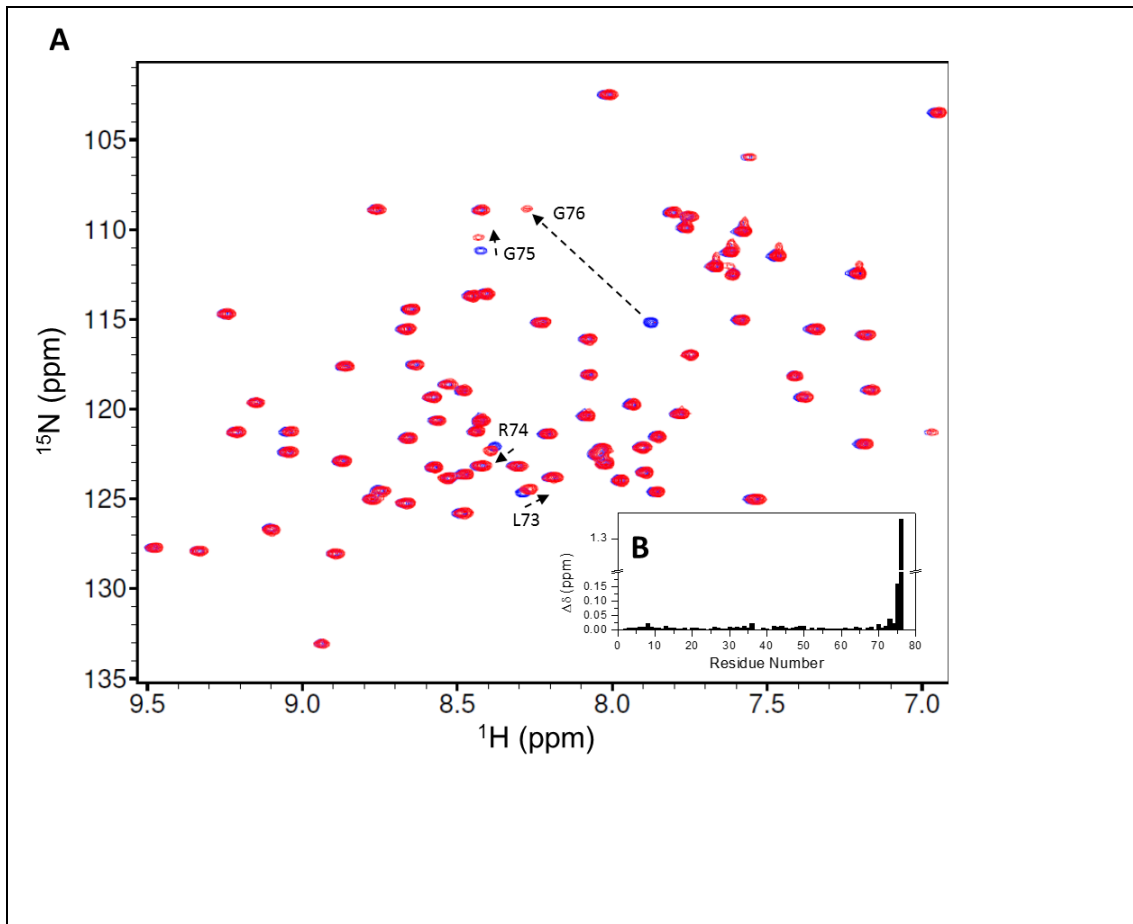
Figure 7. ^{1}H-^{15}N HSQC NMR spectrum of ^{15}N-Ub~(δSH)K77. A: Overlay with ^{15}N-Ub^{WT} (blue); B: Chemical shift perturbations plot (CSP) of ^{15}N-Ub~(δSH)K77.

In order to address the question if the attachment of the mercaptolysine happened through an isopeptide or a peptide bond, the difference in nitrogen labeling between Ub and the (δSH)K residue can be exploited in NMR with a TOCSY experiment (Figure 8A). In its pulse sequence the frequency of ^{15}N nitrogen can be filtered and the correlation between ^{1}H-^{1}H exhibits a tower where the connection of every proton magnetically equivalent. As expected, the mercaptolysine residue has 4 proton signals (Figure 8B). Besides, a shorter mixing time helped to see that the signal corresponding to the epsilon methylene is in fact the closest to the amide bond. Therefore, the preferred amide linkage is an isopetidic rather than a peptidic one.
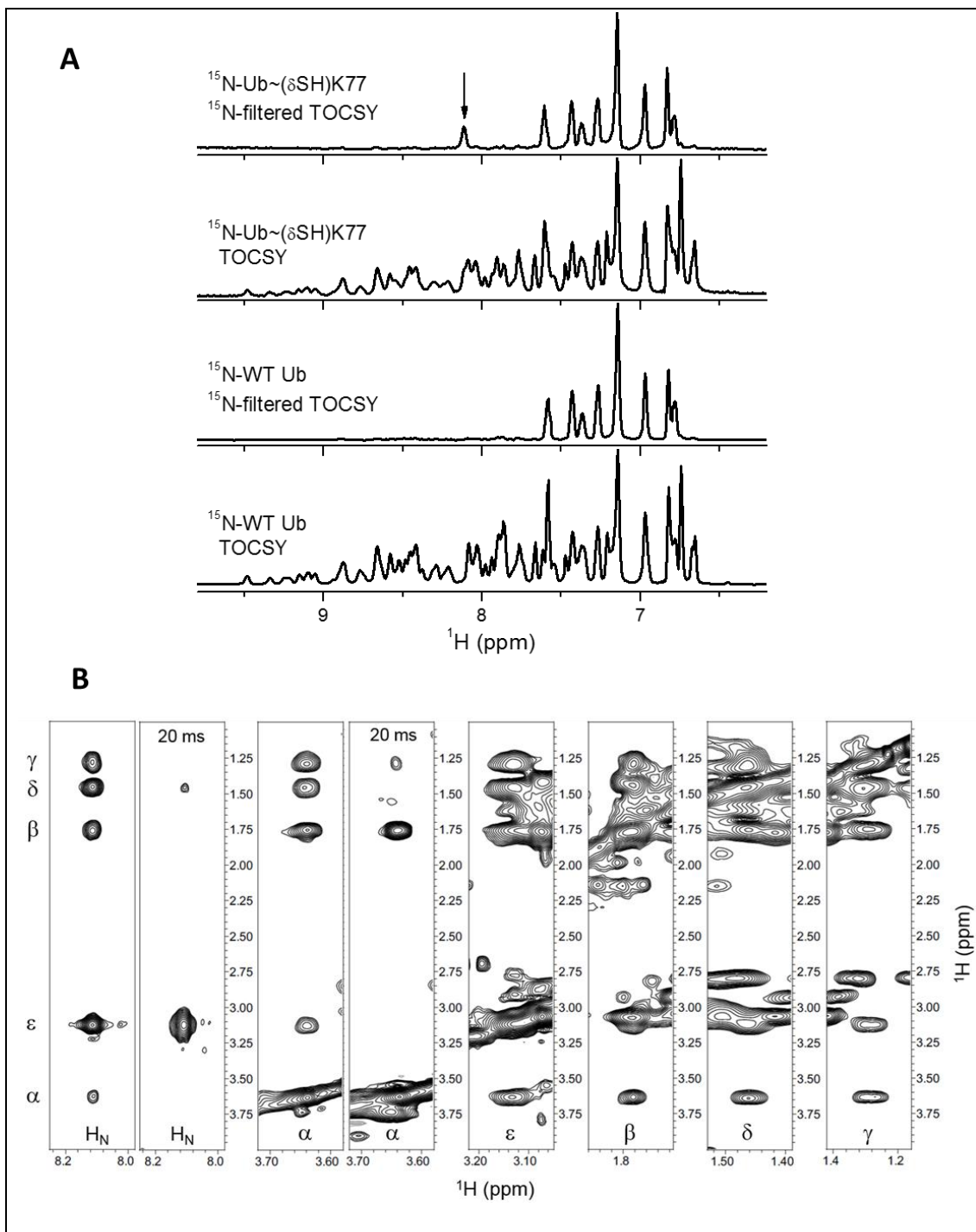
Figure 8. TOCSY ^{1}H-^{1}H NMR spectrum of ^{15}N-Ub~(δSH)K77. A: 1D ^{1}H NMR of Ub~(δSH)K77 from the TOCSY experiment. B: 2D ^{1}H-^{1}H NMR tower slices.

Another direct application of the (SH)K attached to Ub was the structure characterization using paramagnetic relaxation enhancement (PRE). The attachment of an amine oxide as spin label was carried out with an in-house methodology.[54, 55] MTSL (*S*-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl) methyl methanesulfonothioate) was selected to be attached to ^{15}N-Ub~(δSH)K77 as a disulfide between the MTSL and the sulfhydryl present in the lysine derivative (Scheme 22).
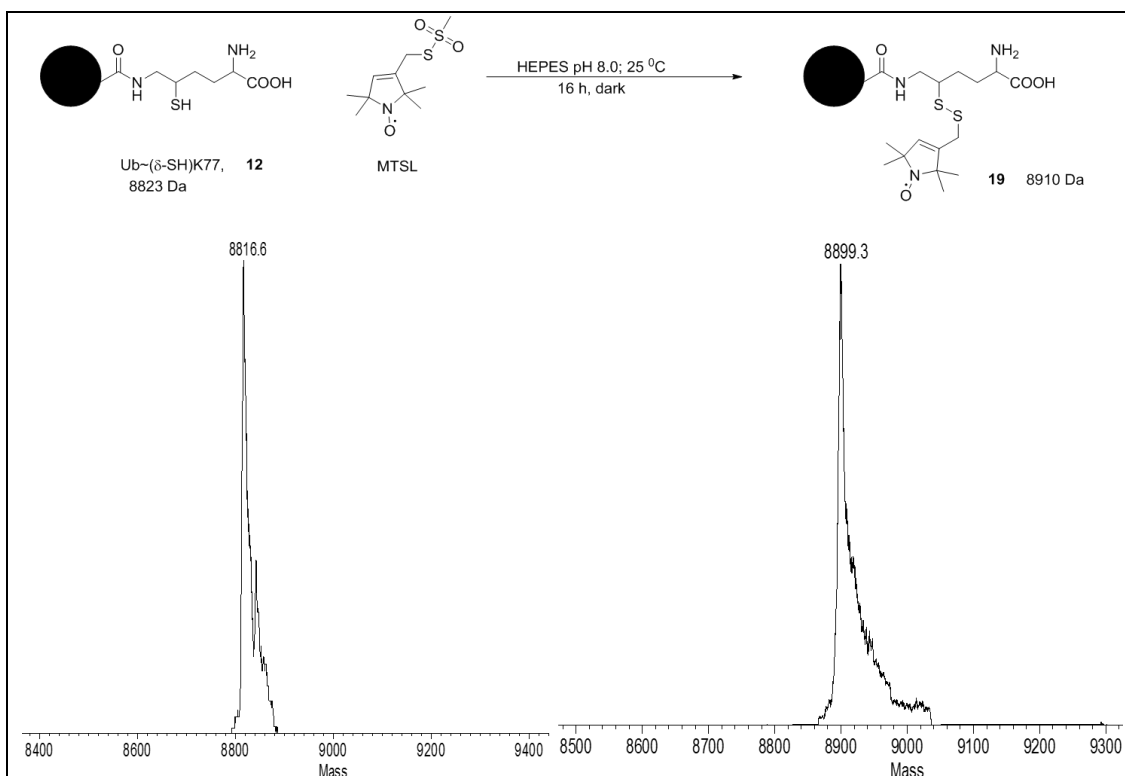
Scheme 22. MTSL attachment to ^{15}N-Ub~(δSH)K77. Reaction scheme and ESI-MS deconvoluted spectra; expected: 8910 Da, found: 8899.3 Da.

The attachment of MTSL resulted in strong attenuation of the NMR signals of the residues in the β1/β2 loop, the β3 strand, and the C terminus of Ub (Figure 9). Moreover, the effect of the paramagnetic effect can be measured when the free radical is quenched with a reducing agent such as ascorbic acid. The quantification allowed the modeling of the spin label position which in this case has an average of 20 Å and approximately 17 Å from G76 (Figure 10) agrees with the expected average location of the C terminus of Ub. Finally, MTSL was removed by reduction with DTT exposing the thiol handle of the mercaptolysine for any other modification.
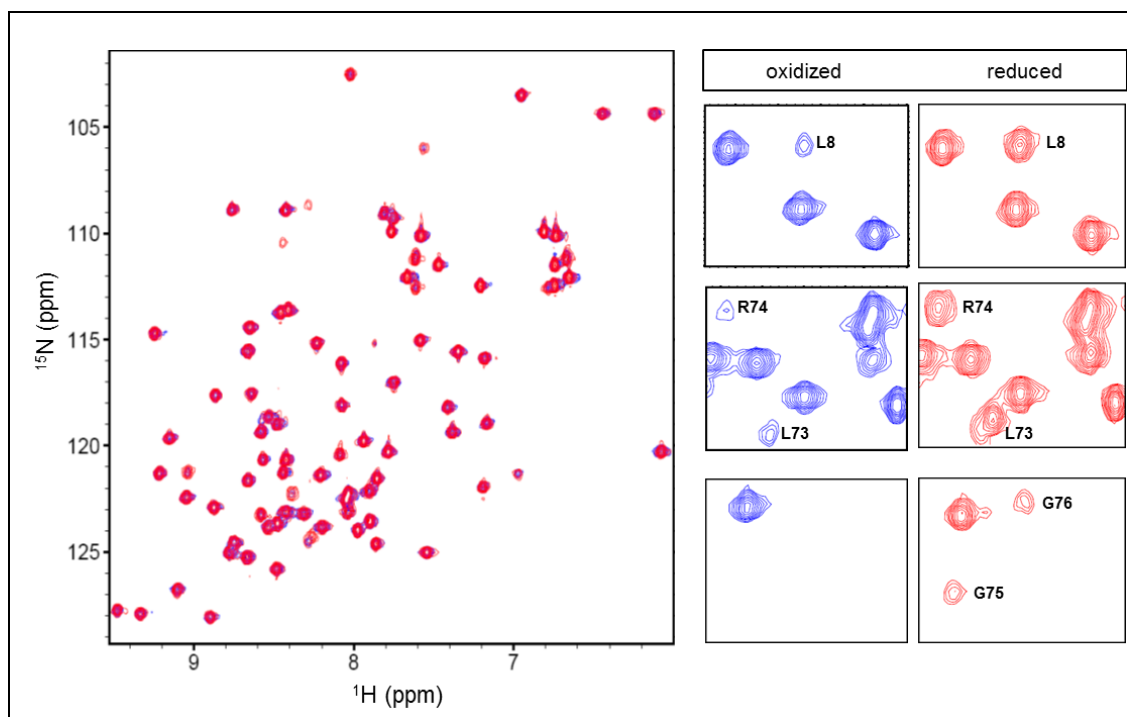
Figure 9. ^{1}H-^{15}N HSQC NMR spectra of ^{15}N-Ub~(δSH)K77 in the presence of MTSL. Blue: ^{15}N-Ub~(δSH)K77, Red: ^{15}N-Ub~(δSH)K77+MTSL; the labels correspond to those residues with strongest attenuations.
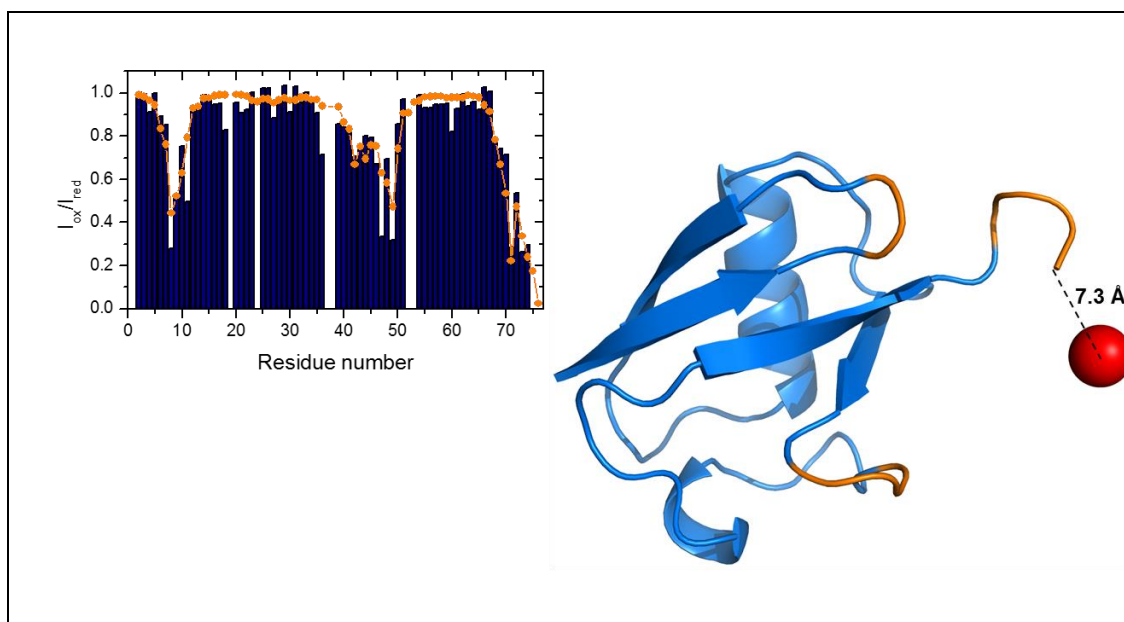
Figure 10. Intensity ratio of ^{15}N-Ub~(δSH)K77 and calculated location of MTSL. Ub is represented as a ribbon model; residues exhibiting attenuations and the C terminus (residues 72-76) are highlighted in orange and the spin label is represented by the red sphere.

In conclusion, alternative and more efficient synthetic routes for the δ- and γ-mercaptolysines **6** and **10** were developed. A thiol moiety as a thioacetate group by a 1,4-Michael addition to a nitro olefin, as well as to an α,β-conjugated aldehyde were introduced. Specifically the thioacetate group as the thiol precursor was chosen because its deprotection is carried out under mild conditions. Moreover, testing the addition of this thioacetate group was successful via free-radical reaction under aqueous conditions using an azo-water-soluble initiator and methanol as co-solvent to keep a homogeneous solution. Finally, NCL methodology was used to elongate Ub at its C terminus with these synthetic mercaptolysines with a yield of 80%, where no Ub$_2$ was isolated. As reported, the NCL RDS still occurs when the ε-NH$_2$ of the mercaptolysines is blocked as a carbamate indicating that the α-NH$_2$ is reactive. The attachment of the mercaptolysines was demonstrated to Ub reversible after incubation with a Ub carboxyl-terminal hydrolase such as YUH1, and no alteration of the Ub

48

structure is detected. Additionally, the utility of an intrinsic thiol handle in the sequence of Ub was used first for the attachment of a fluorophore such as fluorescein-5-maleimide and second to an amine oxide paramagnetic spin labeling where the strongest attenuation effects reside along the hydrophobic patch of Ub, as well as its flexible C terminus was characterized.

Chapter 2: Orthogonality Protection at the ε-NH$_2$ of Lysine for Non-Enzymatic Silver Assisted Ubiquitin Dimerization

As mentioned before, ubiquitination as post-translational modification is controlled by a series of enzymes. In addition, the number of Ub units in the polyUb chain, as well as the lysine residue involved in the Ub-Ub linkage defines the functional outcome of a tagged protein. Delving in the specificity of the isopeptidic signal is an open field for research and the unique conformations a polymer chain can adopt is an intriguing question for understanding and addressing their binding partners.

So far, *in vitro* the successful study of Ub linkages assembled enzymatically requires controlled conditions and the best are achieved when point mutations are made to the polypeptide backbone of Ub. In other words, when lysine residues are replaced permanently, for instance to arginine or cysteine.[56] However, if the E2-E3s are not available, a semisynthetic effort, where the cellular machinery is not used, is critical.

An extra definition must be done here when referring to polymeric chains of Ub. In terms of nomenclature, when working with Ub chains sometimes two concepts are used "interchangeably" to distinguish the monomers from each other. Chemically the monomer in the chain bonded by its C terminus is called the donor and by its position with respect to where the substrate is called the distal domain. Likewise, the monomer in the chain bonded by its ε-NH$_2$ chemically is referred as the acceptor and

by location it is the proximal domain. The nomenclature in terms of distal and proximal domain is extremely useful when any of the monomers in a polyUb chain is ^{15}N-enriched.

In particular a methodology developed by Chin[57] and then adopted but improved in our laboratory[58] offers the compelling combination of synthetic chemistry and site-directed mutagenesis (Scheme 23) within the concept of a convergent strategy. In essence, to achieve the specific isopeptide linkage the donor and the acceptor require being chemoselectively different from other similar groups in the Ub molecule. In other words, they need to be distinguishable and unique. Therefore, the convergent scheme plots the following features:

a) The donor needs all lysines and N-terminus protected.

b) The acceptor needs the targeted lysine exposed while the other amines are protected; stating that the orthogonality in protection/deprotection is an essential synthetic tool.[59]

c) The best protection group for the lysine residues must be cleavable under mild conditions.

d) The donor carboxylic group must be activated, with an easy strategy into a good leaving group.

e) The donor and acceptor must go through a successful chemospecific and chemoselective ligation reaction.

f) The protecting groups must be removed at once also under mild conditions.

To achieve these requirements, the reaction conditions are summarized:

1. The carboxylate of the donor C terminus (distal domain of the dimer) is transformed into a thioester enzymatically with the E1 activating enzyme in the presence of 2-mercaptoethane sulfonate (MesNa) as previously described.[24] Then, all the lysines are protected as carbamates, in particular, with allyloxycarbonyl groups (Alloc).

2. The acceptor ε-NH$_2$ is introduced into the Ub polypeptide sequence by site-directed mutagenesis[50] as an unnatural amino acid[46] (proximal domain of the dimer); and it is introduced as a carbamate, in particular, ε-NHBoc. The other lysine residues in the distal domain are also protected with Alloc groups.

3. The coupling reaction is best performed in basic DMSO, and in the presence of N-hydroxysuccinimide (SuOH) and AgNO$_3$.[60]

4. The mild conditions for deprotection of Alloc groups were achieved using a Ru(II) catalyst in thiophenol, while the deprotection of the Boc group was achieved with diluted TFA.[50]
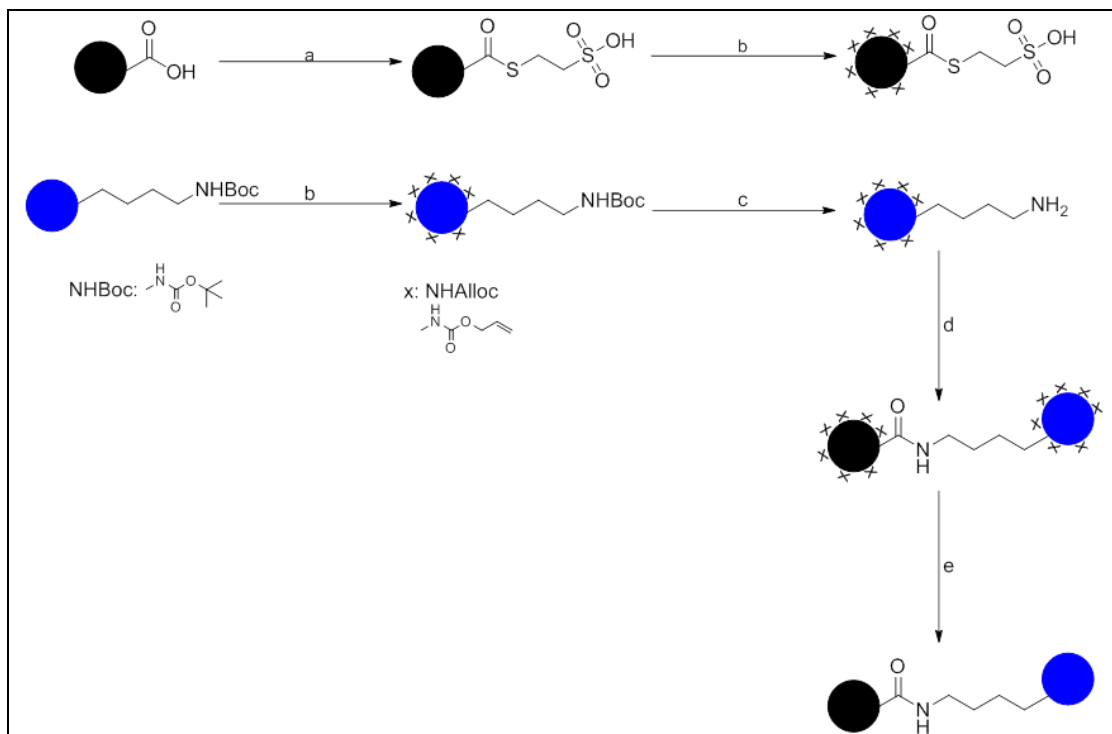
Scheme 23. General scheme for Ub dimer assemblies; a) E1, MesNa, pH 8.0; b) Alloc-OSu, DMSO, DIEA; c) 60% TFA; d) AgNO₃, SuOH, DIEA, DMSO; e) 50%mol [Cp*Ru(cod)Cl], PhSH, DMSO; ~15% yield after 5 steps.

But building Ub chains in this particular way raises some questions due to stability of the protein, and the next four are addressable:

a) Is it possible to exchange the protecting groups keeping the orthogonality intact?

b) Is the reaction yield similar as the original strategy?

c) If the yield is not improved, is there an alternative approach to increase it?

d) Which other conclusions besides a successful coupling of Ub dimers can be brought to light?

In the following sections these questions starting by the exchange of the orthogonal groups will be addressed. The proximal domain would bear an ε-Allyloxycarbamate (Alloc) introduced as an unnatural amino acid; while the other lysine residues would bear a (*tert*)Butoxycarbamate (Boc).

Part A. Carbamate exchange at the proximal domain

In order for the proximal domain to have a ε-Allyloxycarbamate (Alloc) introduced as an unnatural amino acid, the first step of the strategy was the synthesis of N-ε-AllocLys. The experimental conditions previously reported with a quantitative yield of 95% (Figure 11A)[28] and incorporated the N-ε-AllocLys into Ub polypeptide sequence in lieu of Lys 48; referred as K48AllocUb were used. After purification, 10 mg of protein are isolated and lyophilized (Figure 11B).[46] In particular the pyrrolysyl-tRNA synthetase/tRNA pair (PylRS/tRNA) from archaea *Methanosarcina mazei* was used.[61] As mentioned before, an orthogonal pair incorporates into proteins structurally unnatural amino acids. The PylRS/tRNA pair from *M. mazei* responds to the amber stop codon TAG and this system is useful when the recombinant protein is expressed for instance in *E. coli* and the lysine residue bears an ε-carbamate.[46]
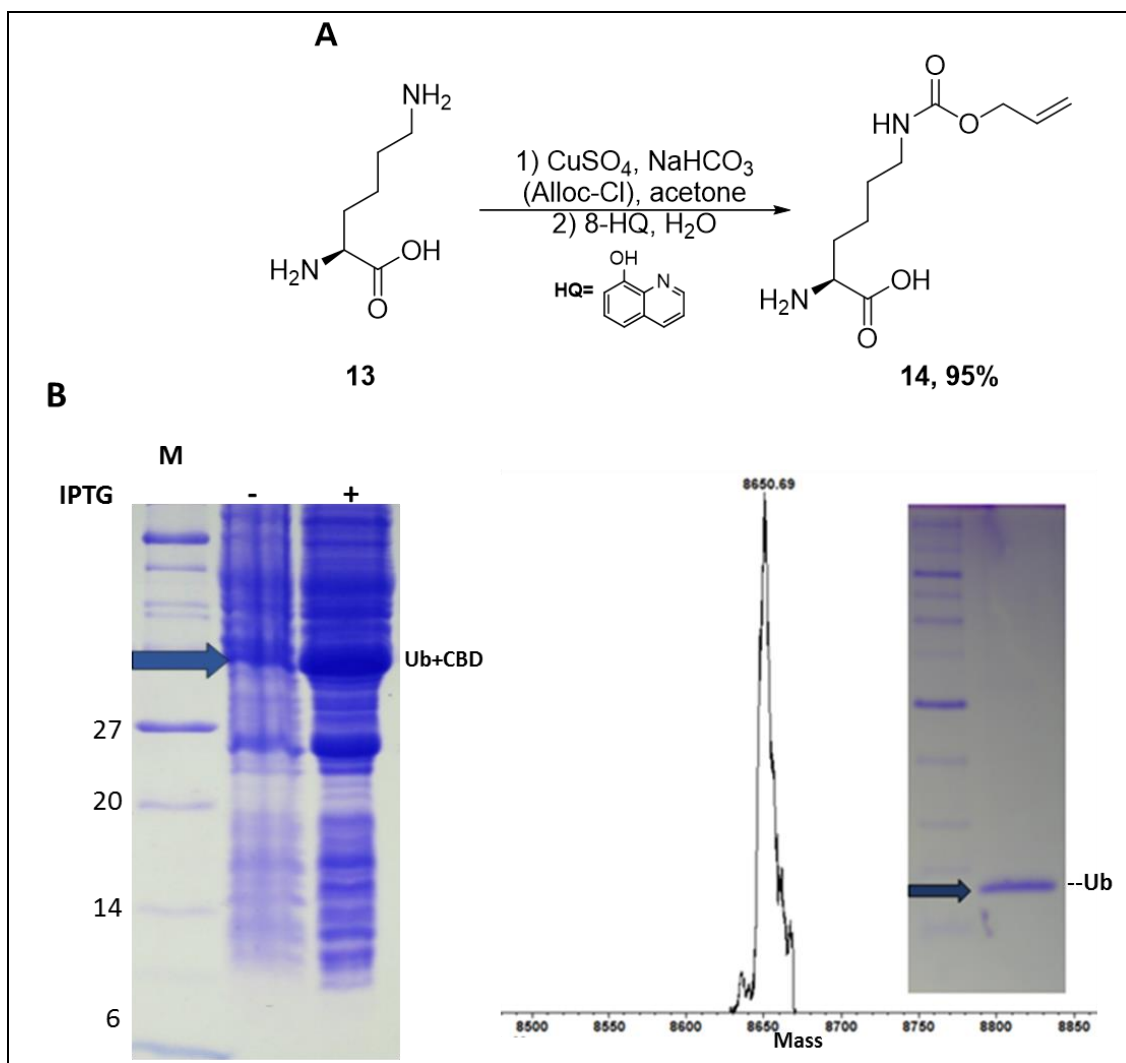
Figure 11. Synthesis and protein expression of K48AllocUb. A: Synthesis of N-ε-AllocLys; B: SDS-PAGE protein expression of K48AllocUb (left); deconvoluted ESI-MS spectrum after chitin cleavage purification (expected 8648 Da, found 8650.69 Da)

The second step consisted in the protection of all the other amines with *tert*-butoxycarbamates (NHBoc). Thus, K48AllocUb was dissolved in basic DMSO followed by the addition of SuOH and Boc anhydride. After ESI-MS characterization, 11 Boc protecting groups are present (Scheme 22). However, the protein is considered fully protected once its mass is higher than 9440 Da (which corresponds to 8 Boc groups: 6 lysines, 1 histidine, and the N-terminus).
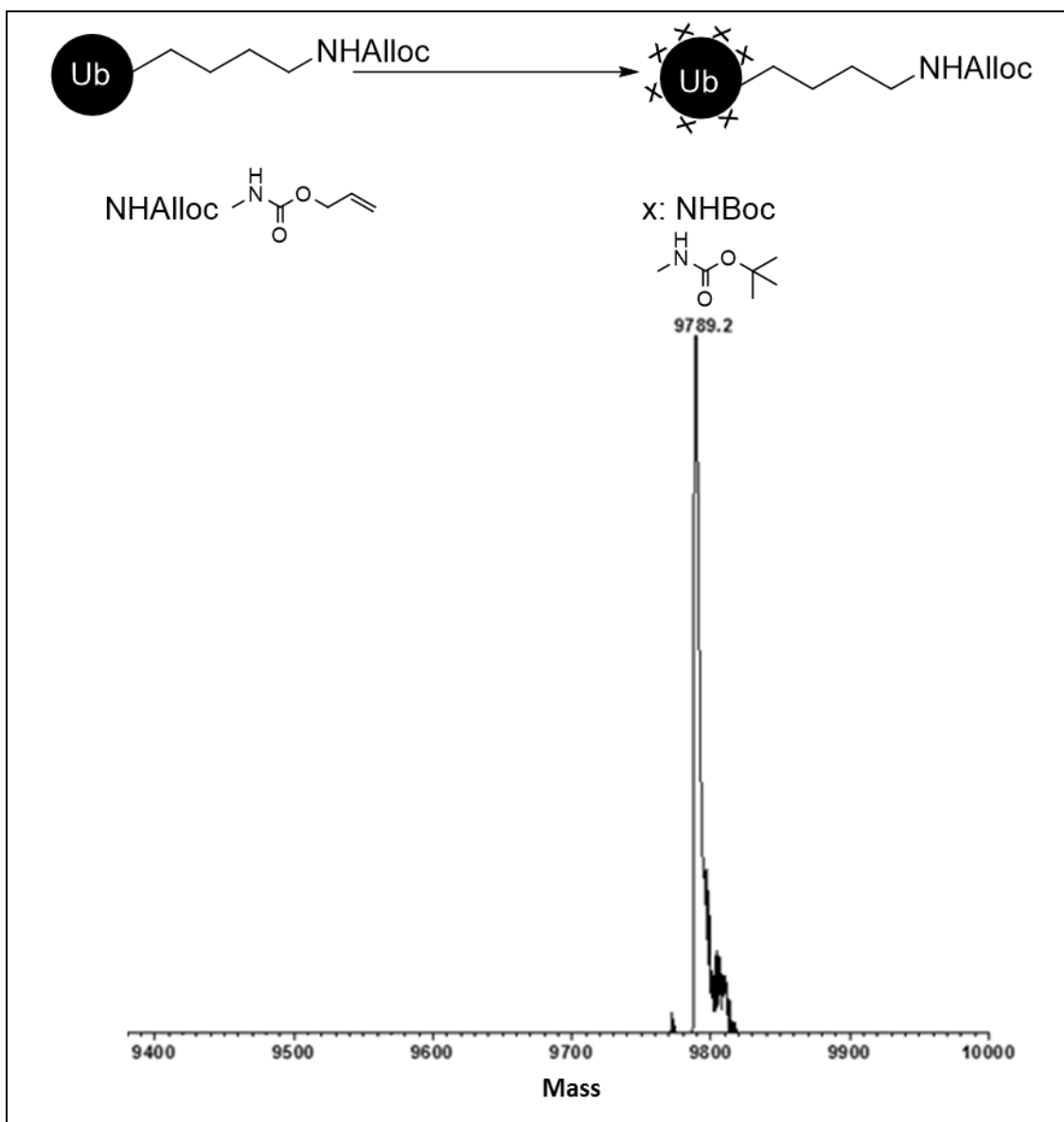
Scheme 24. Lysine side chain orthogonal protection with Boc anhydride. In particular it is shown when 11 Boc groups are present; expected mass 9444 Da, found 9789.2 Da.

All these reactions take place in DMSO, hence require dissolving the protein in DMSO which is miscible with water, ether precipitation must be used to isolate the protein from solution as a pellet. The last step for the carbamate exchange approach was the orthogonal deprotection of the Alloc protecting group. The use of the Ru(II) catalyst is preferred than the traditional palladium phosphine system (Scheme 25).[59]

From the experimental data the orthogonality remained intact because the Boc groups were not hydrolyzed by the Ru(II) catalyst. In terms of yields, starting from 10 mg of K48AllocUb a total of 4 mg of (K48Ub)Boc$_{10}$ were isolated (40%).
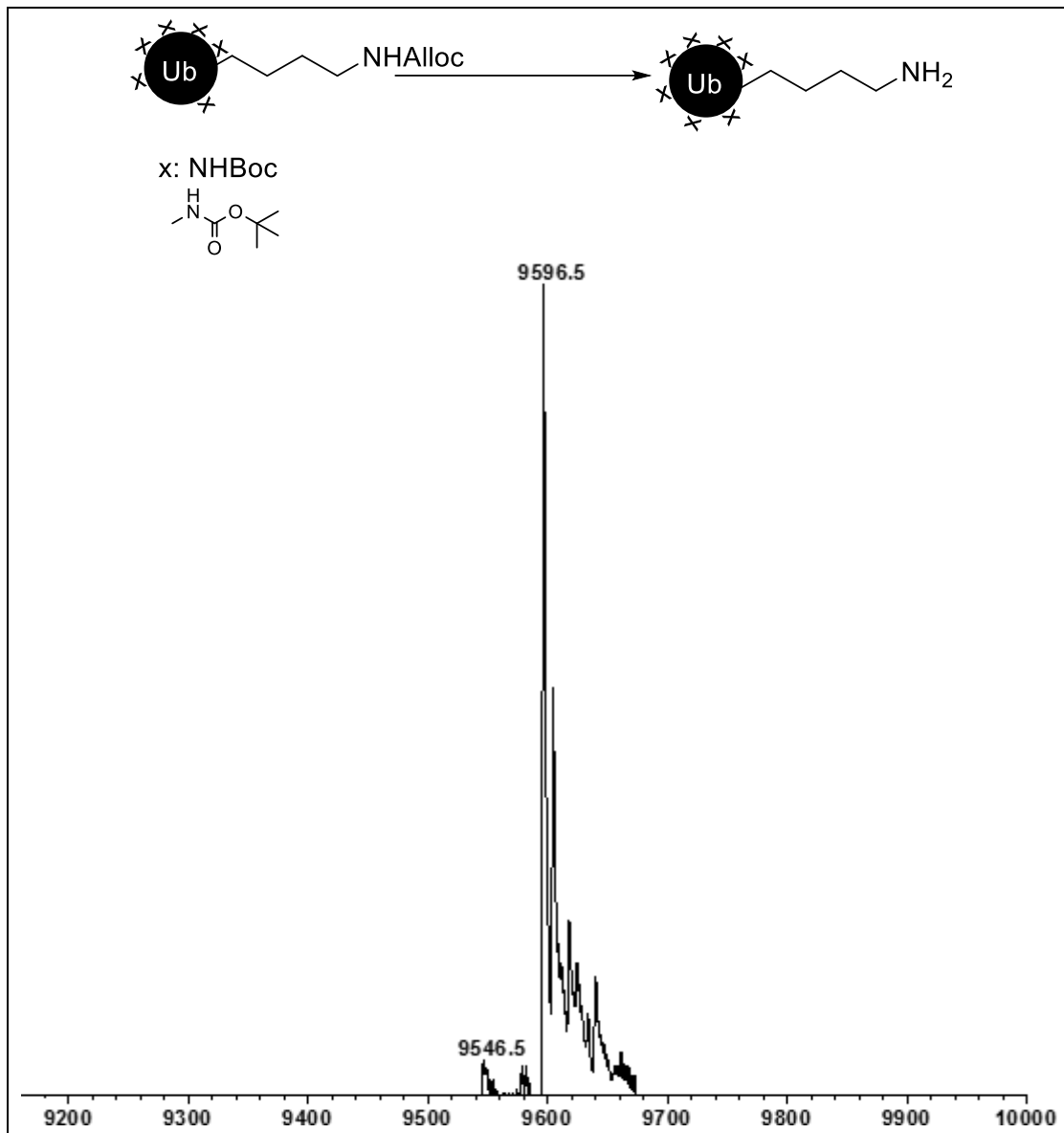
Scheme 25. Orthogonal and chemoselective Alloc deprotection in the presence of [Cp*Ru(cod)Cl]. In particular, it is shown when 10 Boc groups are present (K48Alloc(Boc)$_{10}$Ub: 9681Da; K48(Boc)$_{10}$Ub: 9596Da)

Part B. Silver-assisted ligation

After the donor Ub (distal domain) was activated with a thioester enzymatically,[24] and the lysines were protected with *tert*-butoxycarbamates using the same conditions as in Scheme 24; the ligation reaction was performed as published.[58] After 48 hours of reaction it was noticeable by SDS-PAGE that the reaction yield is lower than 50%, (Figure 12). Increasing the amount of $AgNO_3$, however, did not help to improve the ligation reaction yield. For comparison, the ligation reaction between Ub^{WT} (not protected) with a Boc-distal domain proceeded as expected and, interestingly, no higher polymerization was detected. The last two variables to manipulate were then time and temperature of the reaction. Whatsoever, the yield remained very low < 10 %. Therefore, a strategy where one domain was orthogonally protected with respect to the other domain was followed, as shown in Scheme 26 (C and D). A "mixed ligation" was with the intention of increasing the reaction yield. The idea consisted, as stated before, in having a differential protection in each unit of the dimer before the conjugation reaction. Thus, there are four possibilities to perform the ligation by this approach:

a) Ligation 1 (control): Distal domain as Alloc-protected unit – Proximal domain as Alloc-protected unit.[58]

b) Mixed ligation 1: Distal domain as Alloc-protected unit – Proximal domain as Boc-protected unit.

c) Mixed ligation 2: Distal domain as Boc-protected unit – Proximal domain as Alloc-protected unit.

d) Ligation 2: Distal domain as Boc-protected unit – Proximal domain as Boc-protected unit.
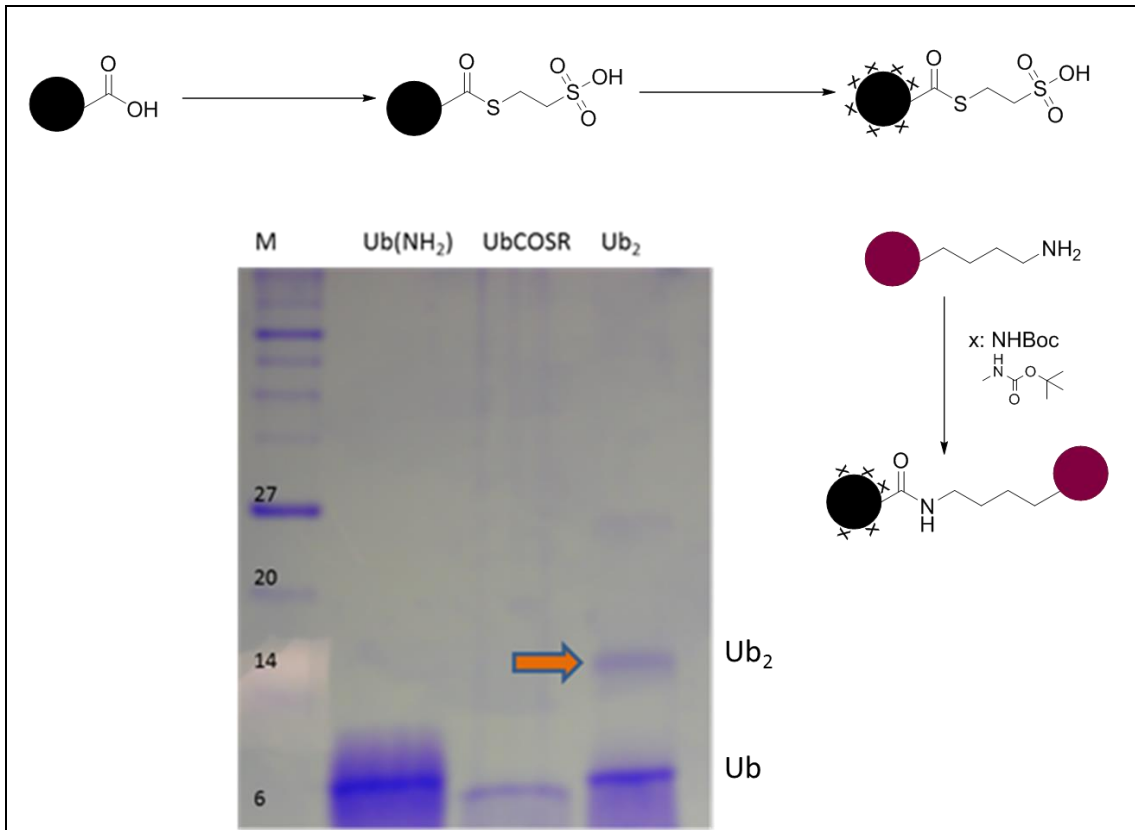
Figure 12. Silver-assisted ligation of K48Ub$_2$; proximal domain Ub distal domain (UbCOSR)Boc$_{10}$
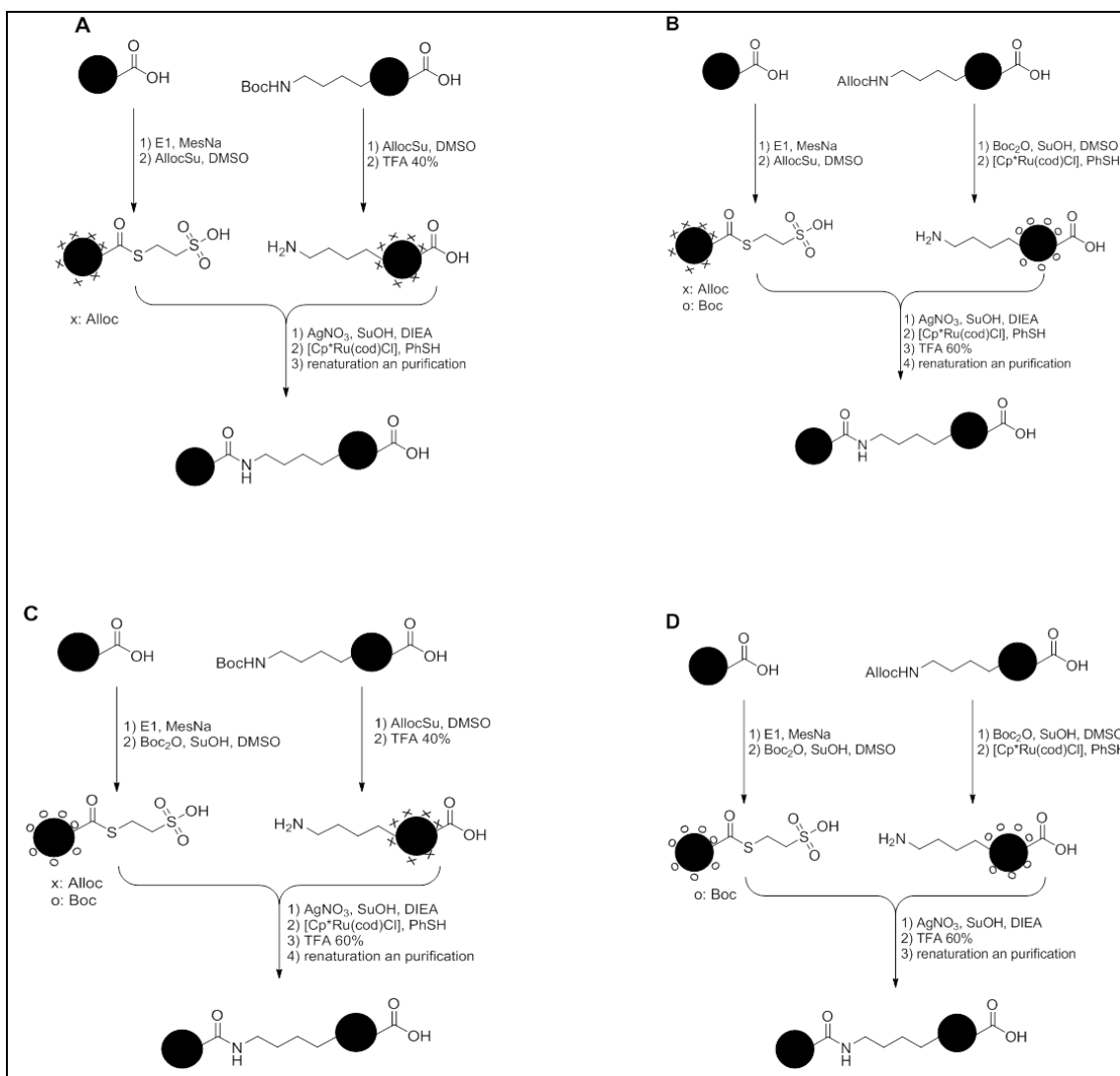
Scheme 26. Control and Mixed ligation reactions of K48Ub. A: Ligation 1 (control);[58] B: Mixed ligation 1; C: Mixed ligation 2; D: Ligation 2.
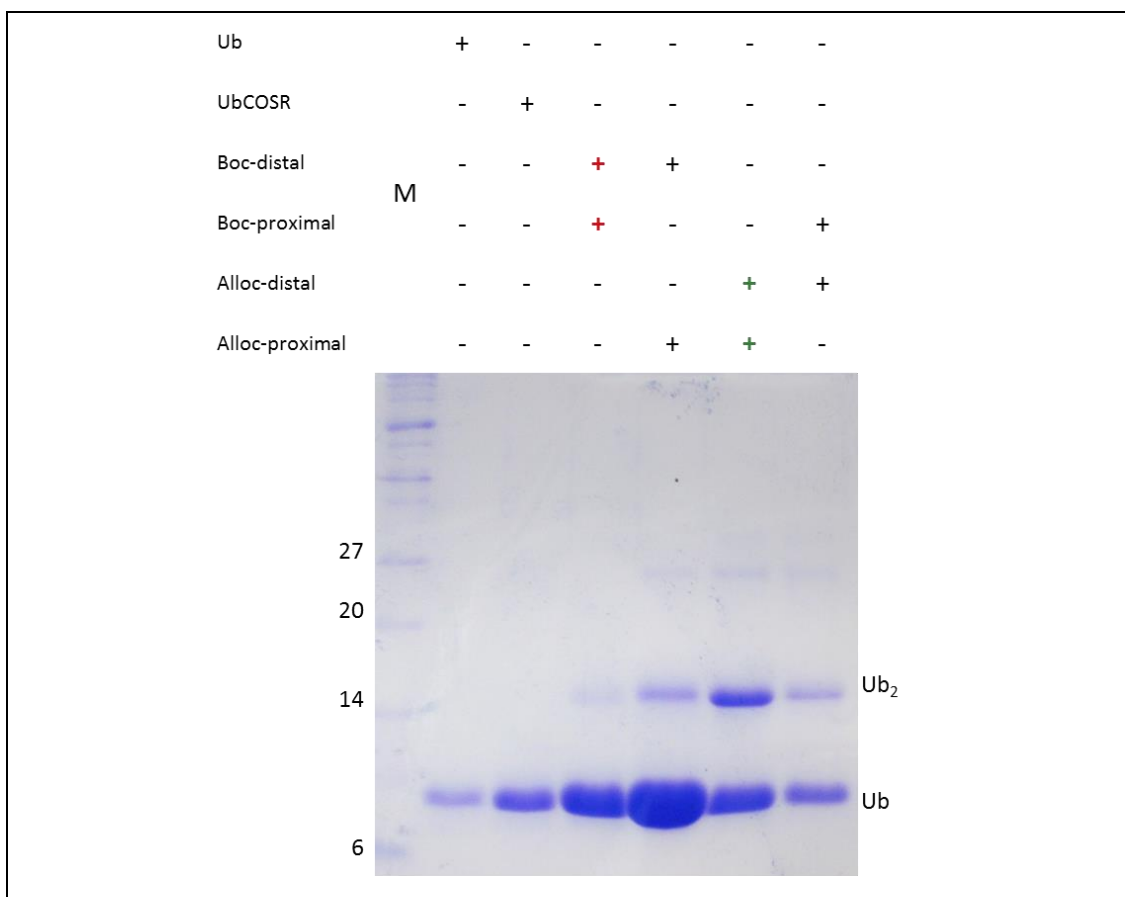
Figure 13. Control and Mixed ligation reactions of K48Ub$_2$ SDS-PAGE after renaturation. Ub and UbCOSR were used as reference. The control reactions with homogenous lysine residues protections are marked with red crosses for *tert*-butoxy carbamates and green crosses for allyloxycarbamates.

Building a chain bypassing the E1-E3 enzymes by the silver-assisted ligation is efficient only under limited conditions. As observed from Figure 13 how prone the monomers are to undergo the ligation reaction suggests the following trend:

$$(K48\text{-}Ub_2)Alloc_{16} \geq (K48\text{-}Ub_2)Alloc_8Boc_8 \approx (K48\text{-}Ub_2)Boc_8Alloc_8 \geq (K48\text{-}Ub_2)Boc_{16}$$
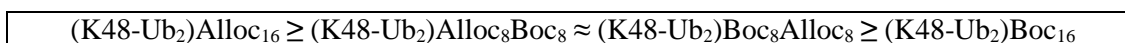
Moreover, at this stage, after deprotection and purification of the dimers by cation exchange chromatography from 10 mg of monomers (5 mg each); 2 mg of dimer are isolated for the best condition of reaction,[58] in other words, the yield drops to a 10%. And for the mixed ligation, the yield is below 5% (1 mg). Therefore, answering the question if the reaction yield is similar as for the original strategy, and if this could an alternative approach to increase it, at least in my hands the answer is negative.

Part C. Other related conclusions

One of the reasons why the strategy is not successful, when Boc groups are present, might be the ligation reaction as the crucial step of the methodology. The ligation relies in the fact that donor and acceptor groups must be in closer proximity for the silver-mediated intermediate to occur;[60] and even when the polypeptide chain is totally unfolded, the presence of tert-butoxyl groups might create steric hindrance for that close proximity to happen. A proposed mechanism of this ligation is depicted in Scheme 27. The thioester of the distal domain needs to be activated by the silver cation. To stabilize the silver chelate it is possible the base also participates donating briefly its electron pair as discussed previously.[60] Then, the N-hydroxysuccinimide can make the nucleophilic attack on the carbon of the carboxylate. As a consequence the esterification of the C terminus is performed. The product of this esterification is a succinimide ester, which is a better leaving group than the previous thioester. Therefore, the primary amine of the proximal domain can react with the ester and the ligation reaction concludes with the formation of an amide bond.
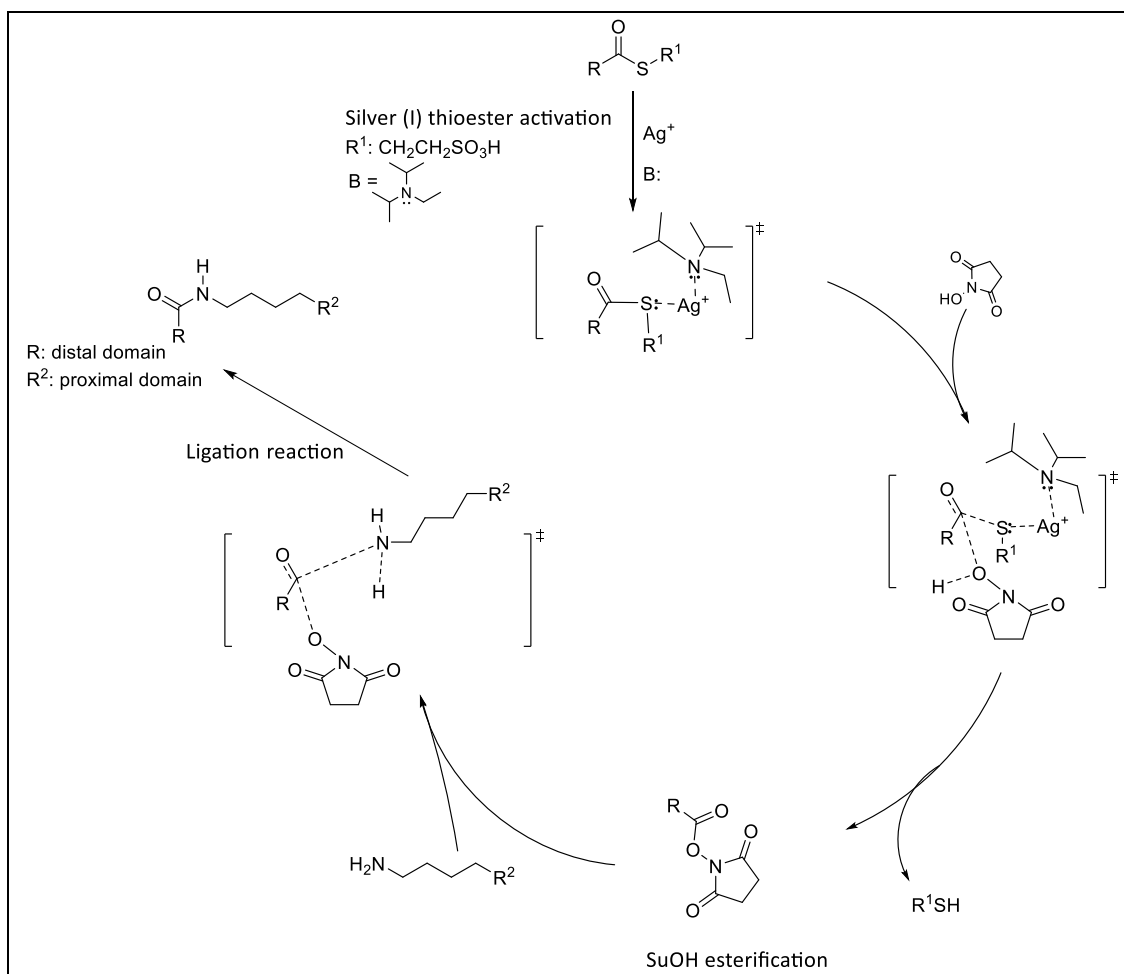
Scheme 27. Proposed silver-assisted ligation reaction mechanism

In terms of experimental variables one of the first challenges is the solubility of the protein in DMSO, which was chosen because it is a polar aprotic solvent and not nucleophilic. Fortunately Ub is partially an easy protein to dissolve in DMSO and to avoid working in not homogeneous it needs to be incubated for longer periods of time before increasing the pH. Next, extensive ether precipitation rounds, in order to get a solid precipitate, caused the reaction yield to decrease. For the ligation conditions, when the temperature or the concentration of silver nitrate was increased, the yield was not boosted. In addition, a closed monitoring of the reaction depend upon its detection by mass spectrometry, higher resolution mass spectrometry is

63

required since the protein do not have any positive charge in its polypeptide chain resulting in a not so easy way to detect it at least by a time of flight (TOF) detector.

However, it is interesting to notice the use of the Ru (II) catalyst for the Alloc deprotection. Previous studies showed the flexibility of the catalyst which survives air and water exposure.[62] It is noticeable the catalyst depends on a sulfur source, which in comparison with the Pd (II) catalyst will result in poisoning the metallic center.[59] Therefore, combining what is known about Ru (II) and Pd (II) allyl-deprotection[63] the catalytic cycle depicted in Scheme 28 is proposed.
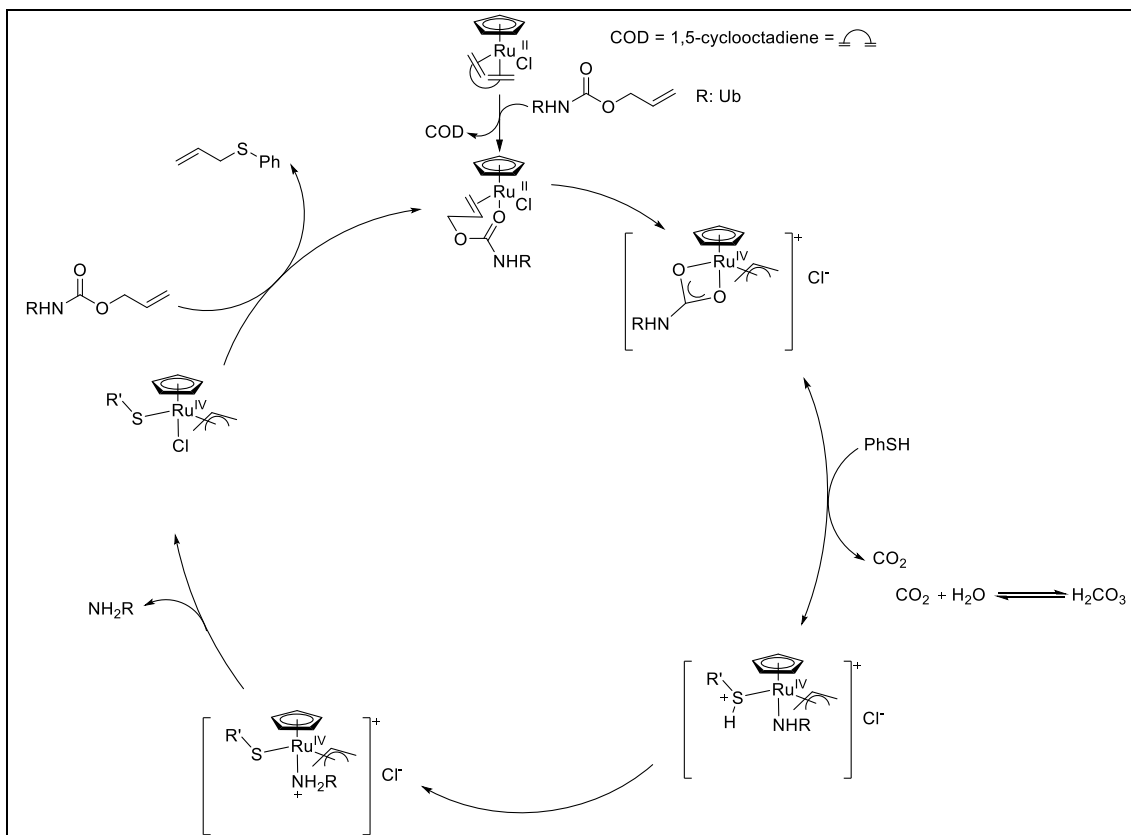
Scheme 28. Catalytic cycle cleavage of allylcarbamate by Ru(II).

64

In the catalytic cycle the allylcarbamate coordinates with the Ru (II) center displacing the cyclooctadiene ligand. By a C-O bond cleavage, the allyl group coordinates with the metallic center changing its valence status from Ru (II) to Ru (IV). The allylic system must contribute with $4\bar{e}$ by coordination and not by retro donation; as a consequence, the metal reaches its stable electronic configuration with $18\bar{e}$ even though its oxidation state is the highest. The coordination of the thiophenol (sulfur source) in the cycle displaces CO_2 but keeps the Ru center at its higher oxidation state. Moreover, the thiol does not poison the organometallic catalyst. In order for the amine to be released from the metallic center, an acid-base reaction takes place. The ammonium salt is then displaced by the chloride ion when another molecule of carbamate coordinates with the metallic center. A C-S bond between the thiol and the allyl group helps the cycle to restart. Therefore, the metallic center is reduced to Ru (II). The fact that a higher concentration of [Cp*Ru(cod)]Cl is used and the temperature is higher than room temperature, gives the idea that the deprotection is not favored kinetically. Though keeping the orthogonality remains intact. If Pd(II) were used instead, when the sulfur source coordinate with it, its lifetime would remain shorter and the number of turnover cycles will be diminished resulting in an inert catalyst over a shorter period of time.

In conclusion, the non-enzymatic ligation strategy when the orthogonality is exchanged even though it was thought to be milder in terms of experimental conditions is not as successful as the one already developed. Moreover, it remained as a constant challenge to expose other type of proteins to conditions in which their

stability might be compromised, in terms of refolding and return to their tertiary structure; when ubiquitination has to be tested. Achieving the K48-linked Ub dimer in order to compare it with an enzymatically synthesized dimer resulted in a sample quality not ideal for measuring it by NMR. Therefore, another approach for dimerization of proteins was targeted.

Chapter 3: Lysine 48-linked Ubiquitin Dimers through Isopeptide Bond Mimics

As mentioned several times, when a protein is tagged for ubiqutination a specific E2-E3 set is needed and even though the K48Ub-linkage is the most common signal for degradation, it is not the only factor in this process. UBDs and DUBs are also needed for the process to take place in a regulated manner and to keep the Ub concentration constant. The formation of the isopeptide bond between the protein and Ub tag becomes crucial and the conformation, for instance of the complex between Ub and the E2-E3 system or Ub-UBDs, or Ub-DUBs are very interesting systems of study which would aid to design better inhibitors if a therapeutic focus is on drug discovery to benefit patients with these neurodegenerative disorders.

A dimer of Ub linked through an isopeptide bond by its K48 residue adopts a closed conformation when it is near the physiological pH.[55] Those UBDs which recognize this particular dimer conformation interact with the surface created by those hydrophobic patches.[64] Thus, following the trend of bypassing the E1-E3 system, other semisynthetic strategies had emerged. In particular, the most relevant are those where the chemical reactions can be selected to mimic the isopeptidic bond keeping the properties of being a dynamic linker, in terms of having rotational degrees of freedom and conformational changes. Therefore, ideal systems which simulate a dynamic linker for UBD interactions vary from very rigid (fixed)

structures or closer to the natural linkage. The following examples are based in modifying the C terminus of Ub either by removing G76 and deal with Ub75 thioester species in order to build the isopeptide linkage; or, building the isopeptide bond by coupling a lysine to cysteine mutation in the presence of an electrophilic functional group.

Part A. Representative examples of semisynthetic dimerization

    I.   Copper-assisted click chemistry

The isopeptide bond was built as a triazole by Cu(I) azide-alkyne cycloaddition reaction for UBD binding pattern.[65] The azide (p-azido phenylalanine), semi-precursor of the acceptor was incorporated by site-specific unnatural amino acid mutagenesis.[66] The alkyne, semi-precursor of the donor was introduced by intein chemistry (Scheme 29).[48] The aromatic isopeptide surrogate was used to establish if the UBA domain of Mud1 protein would have the same affinity as the natural dimer which exhibits a K_D ~1 µM, and it is known by NMR to bind to K48Ub$_2$ in a "sandwich-like" manner.[67]
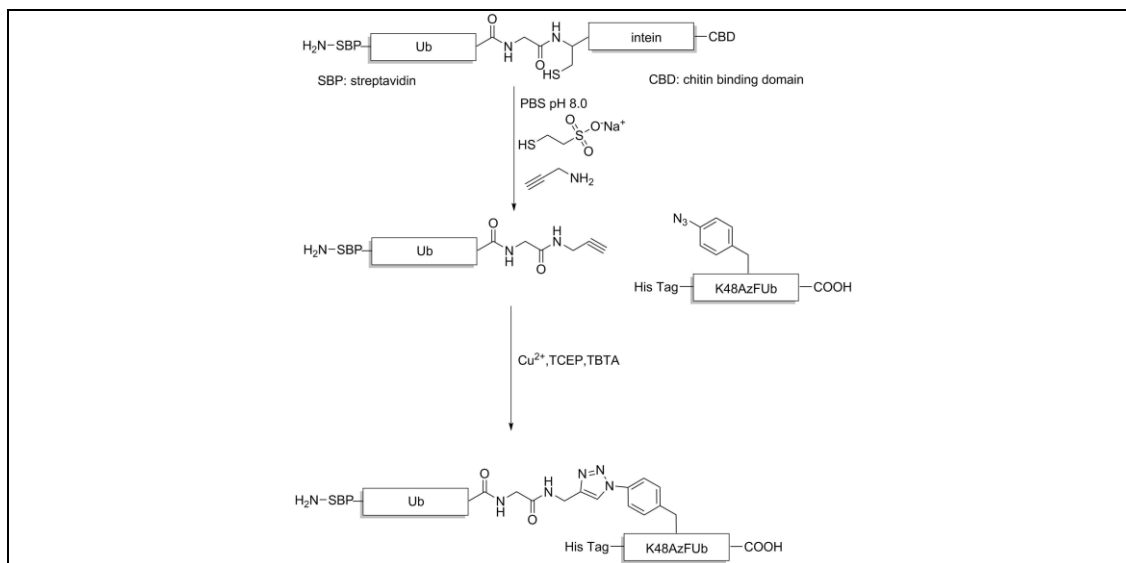
Scheme 29. Click chemistry isopeptide surrogate

II. Michael acceptors

a) Linker size

Ub-Isopeptide dimer probes were built by SPPS where the acceptor Ub bears a vinylamine fuctional group which is coupled with the donor G75Ub thioester in the presence of SuOH.[68] In order for the acceptor to have the vinylamine system, glycinol was oxidized by the Swern reaction and the vinyl system was introduced via the Horner-Wadsworth-Emmons reaction. Then, the vinyl system was used to react with different DUB enzymes; showing that the isopeptide bond mimic is a DUB scavenger since the product of reaction is a C-S bond; therefore an increment in MW is detected by SDS-PAGE (Scheme 30).
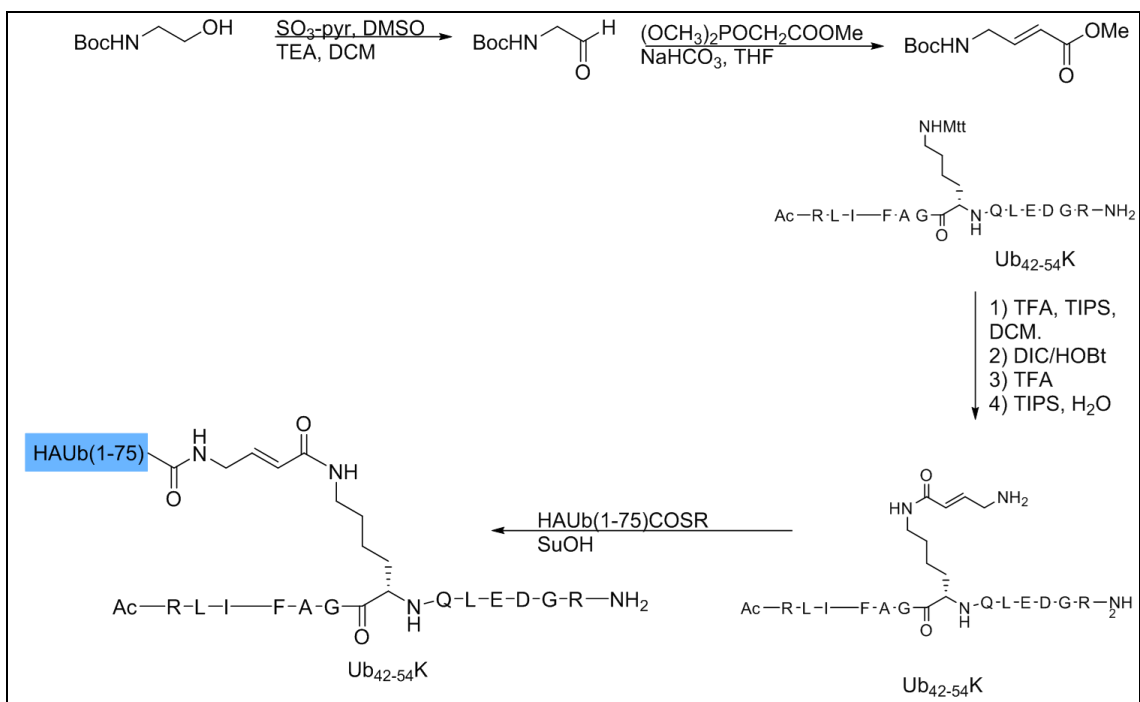
Scheme 30. Ub-isopeptide probes by vinylamine

b) Dichloroketone linker.

Nonhydrolyzable analogs of Ub dimers towards inhibiting deubiquitinating enzymes as stable systems were prepared in order to understand the enzymatic specificity, affinities, and binding sites.[69] In particular, to build the dimer the donor and the acceptor Ub were mutated as cysteine residues. To build the isopeptidic bond, dichloroacetone was used as the electrophilic system at slight alkaline pH conditions (Scheme 31). Later, a kinetic DUB assay was performed using these synthetic Ub dimers; with this assay the activity of the different dimers towards deubiquitination showed the level of inhibition. In summary, DUBs UCH-L3 and USP5 were used to measure the level of inhibition. UCH-L3 was inhibited by $^{11\text{-}76}Ub_2$, $^{29\text{-}76}Ub_2$ and $^{48\text{-}76}Ub_2$ since they are close enough to the enzyme surface. USP5 was in contrast, greatly inhibited by $^{29\text{-}76}Ub_2$.
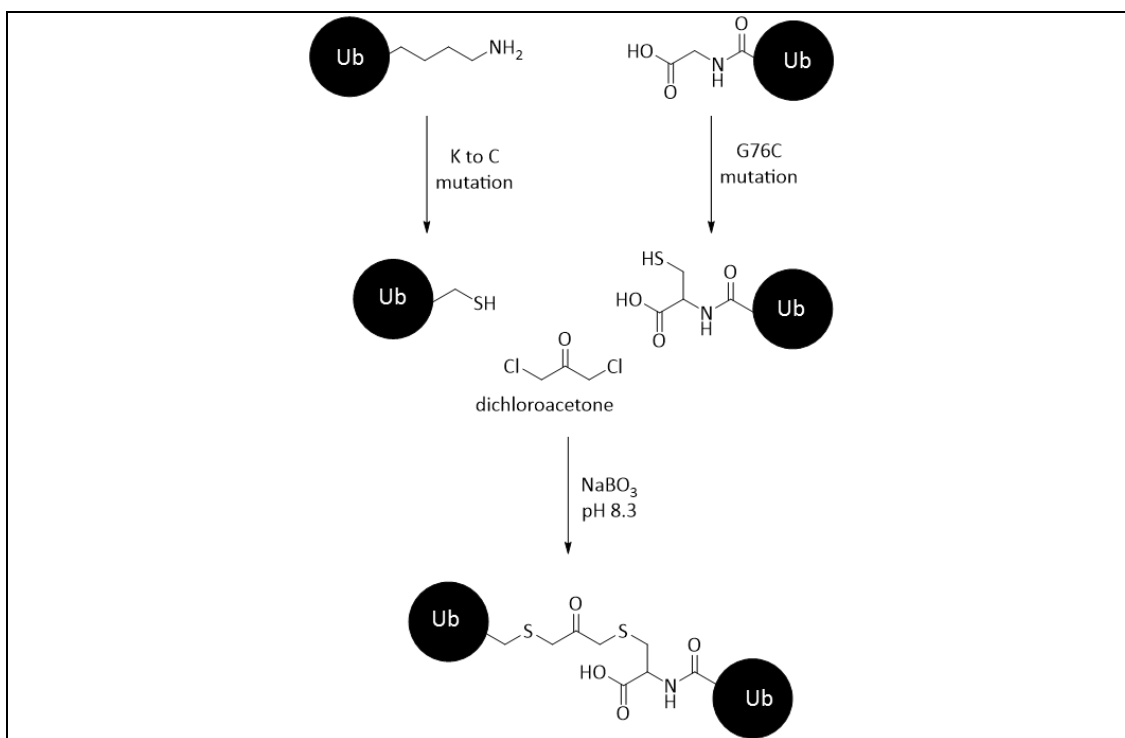
Scheme 31. Synthesis of lysine surrogates with dichloroacetone.

III. Thiol-ene chemistry

A photochemical approach to produce a C-S bond in the isopeptide bond was used by taking advantage of the addition of thiyl radicals to alkenes.[70] Without side reactions or oxidation of any methionine present in Ub, the donor monomer was modified as an allylic system whereas the acceptor monomer was a cysteine mutant. Then, in the presence of a synthetic phosphinate (in particular, lithium acyl phosphinate) the monomers were coupled (Scheme 32).[71] To test the applicability some DUB enzymes were used to determine if the isopeptide mimic was cleaved.
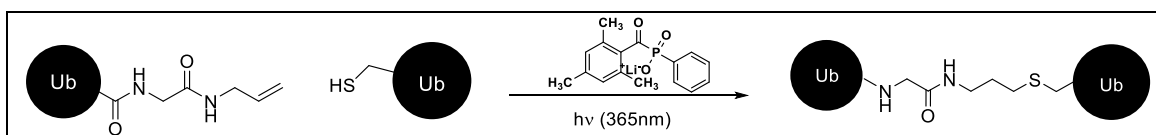
Scheme 32. Photochemical thiol-ene click chemistry

IV. Intein-Disulfide exchange chemical ubiquitination.

By the use of cysteamine the C terminus of Ub was modified to introduce a thiol moiety.[72, 73] Then, it was treated with the Ellman's reagent, DNTB to have a stable but cleavable disulfide. At this stage PCNA[72, 74] or H2B[73, 75] known lysine residues were mutated to cysteine residues and used for the coupling reaction at slightly alkaline pH, as shown in Scheme 33.
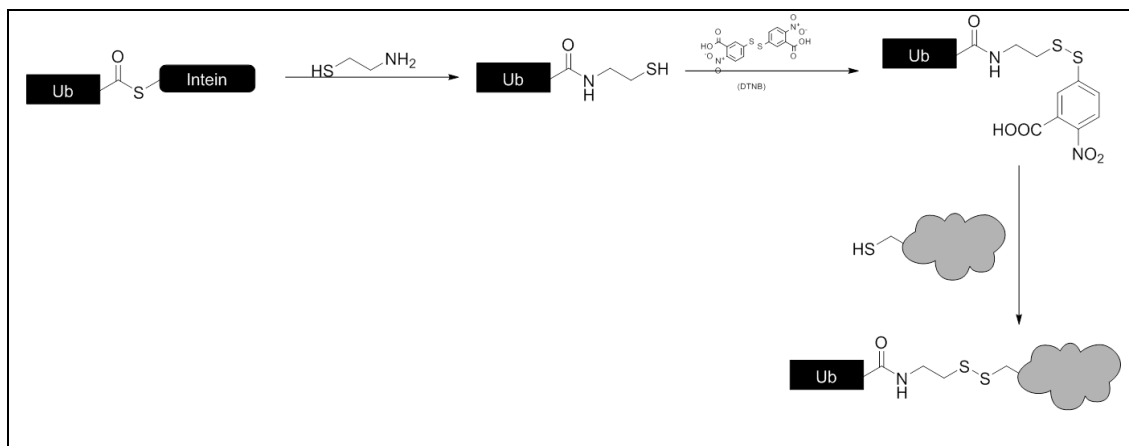
Scheme 33. Disulfide exchange chemical ubiquitination

These two last approaches, in particular, are quite simple strategies though the isopeptide bond is not entirely natural. This means the linkage is one bond longer and at least one sp^3 Carbon atom was substituted by Sulfur. Centralizing the attention in the K48-Ub$_2$ systems the approaches shown in examples III and IV were selected as methodologies for exploring water-soluble azo-compounds, focusing the attention on answering the question if these isopeptide bond mimics would exhibit the same pH-dependent conformational switch as the natural dimers. In the case of the thiol-ene click chemistry, the experimental conditions for producing thyil radicals under thermal decomposition of the initiator were optimized for V-50. In parallel, for the disulfide exchange diamide was used as the redox azo-compound, which so far it has only been used to oxidize glutathione for *in vivo* studies.[76] In order to answer the pH dependence, NMR spectroscopy was the characterization technique to assess if the hydrophobic patches are or not in close proximity. All systems were compared with the homologous natural dimer (built enzymatically).

Part B. Syntheses of K48Ub₂

a) Natural dimer

For building the K48Ub2, the controlled enzymatic methodology previously reported was followed.[56] To control the polymerization reaction the donor was blocked as a K48-Boc lysine through site-directed mutagenesis.[50] Likewise, the proximal domain was blocked with an aspartic residue (Scheme 34). After the incubation with E1 and E2-25K (K48-linked specific conjugating enzyme), the carbamate and aspartic residue were removed with diluted TFA and YUH1, respectively. From 10 mg of distal monomer, approximately 15 mg of dimer were isolated after cation purification (75% yield). In order to study the dimer easily by NMR each domain was ^{15}N isotopically labeled and ligated with its not isotopically enriched counterpart separately.
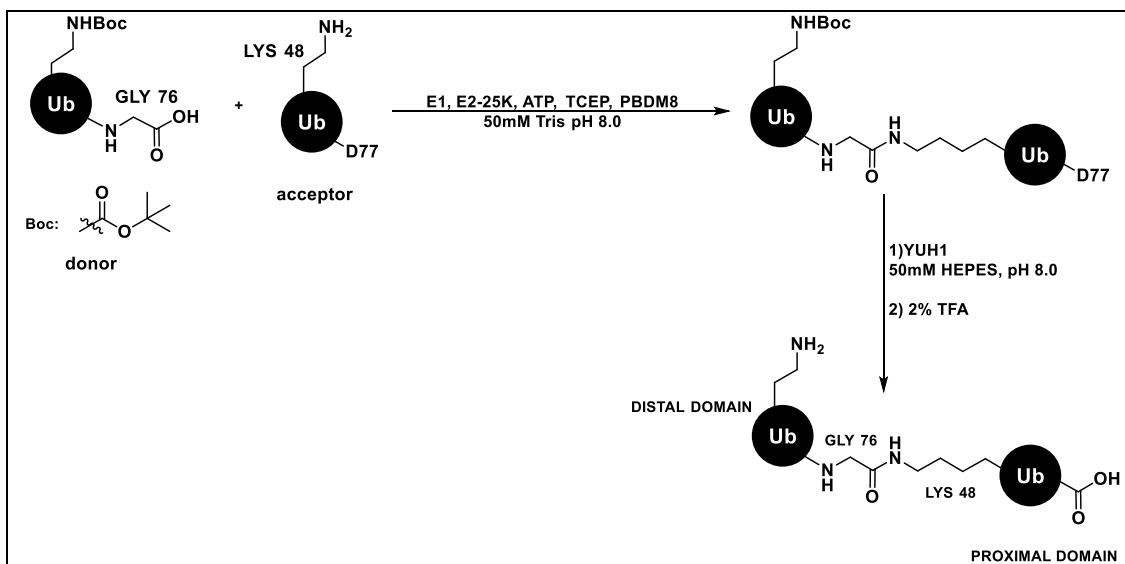
Scheme 34. Synthesis of K48Ub₂ natural dimer

b) Thermal thiol-ene click chemistry

As previously reported by Strieter's research group the acceptor Ub monomer was the recombinant Lys to Cys variant of Ub, K48C Ub; and the donor Ub monomer was modified as an allylamine using D77Ub as the starting material and YUH1.[70] For easy terminology, Ub-allylamine will be referred as the previous reports, UbAA. It is important to notice that after two hours of incubation with the enzyme Ub-AA was also hydrolyzed to Ub, lowering the yield from 80% to 40%. However, since Ub can be recovered after cation purification, Ub-AA could be also produced from the direct incubation at pH 8.0 of UbCOSR (as shown in Scheme 35). As a consequence a convergent synthesis of the donor is achieved keeping the yield between 80-90%.
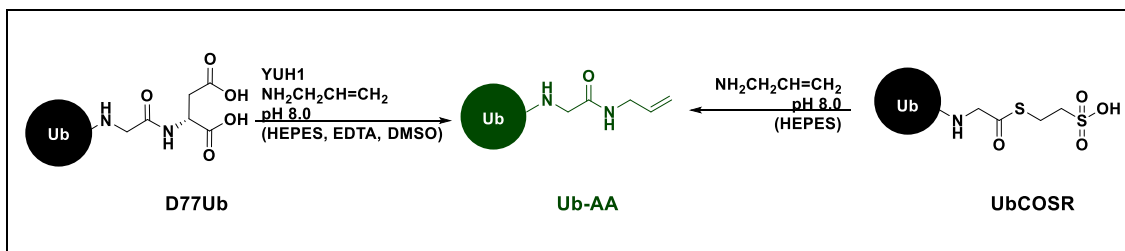
Scheme 35. Convergent synthesis of Ub-allylamine

By the thermal activation of the V-50 initiator the thiol-ene chemistry reaction was performed without special photochemical cells or reactors under more accessible conditions. In fact, when Strieter's group scanned their customized photochemical initiator, as a control they used the same thermal initiator.[70] However, the concentration used of V-50 was 100 molar equivalents and the sole product was the disulfide dimer, K48C Ub_2. Therefore, the V-50 concentration was screened and the reaction conditions needed to minimize the disulfide dimer and favor the outcome of the reaction toward the thiol-ene chemistry product of reaction (Scheme 36A) were

optimized. First, at its half-life decomposition temperature and time (50 °C, 10h) the reaction was monitored by SDS-PAGE under non-reducing conditions (Scheme 36B). The selected concentration of the thermal initiator was 5 mM. However, the yield of the reaction is not particularly high since some amount of monomer is still present as observed by SDS-PAGE. Thus, the second variable screened was the temperature. I tested the reaction at 37 °C in the presence and absence of β-mercaptoethanol (βMe) as reducing agent (Scheme 37). Under these conditions the thiol-ene K48Ub$_2$ C-S was isolated in higher yield (50%) once the dimer was purified by high performance cation exchange chromatography from UbAA. The profile of elution was confirmed by ESI-MS.
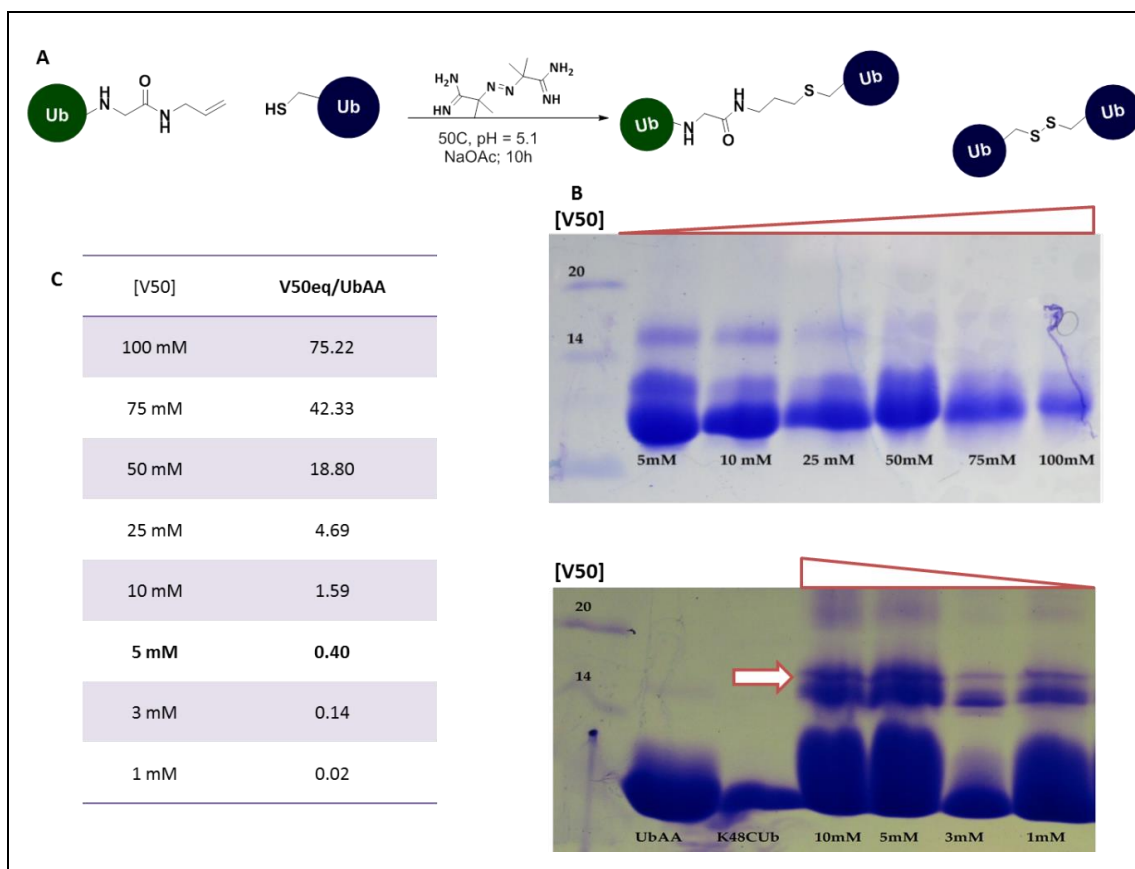
Scheme 36. V-50 concentration screening; A: Reaction strategy; B: SDS-PAGE of V-50 concentration scanning; C: Comprehensive table of equivalents used in the reaction
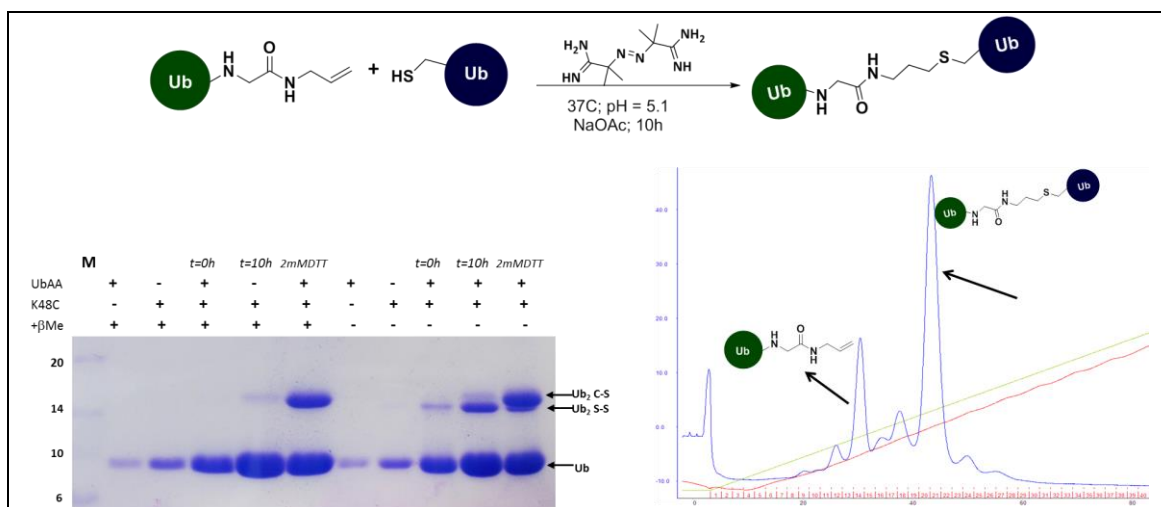
Scheme 37. Thermal thiol-ene chemistry at 37°C; Reaction strategy, SDS-PAGE and cation chromatogram

c) Intermolecular disulfide exchange

To build the disulfide K48-linked diubiquitin isopeptide bond mimic the first part of the methodology reported by Muir, *et. al.*[73, 75] and also used by Zhuang, *et. al.*[72, 77] was implemented. However, instead of using DTNB as the oxidizing agent for the ligation of the monomers, diamide, a water-soluble azo-redox molecule was used as the oxidizing agent. Kosower's review about diamide was extremely useful because it described the compound in extensive detail, stating its stability and storage, mechanism of reaction with thiols and its intermediates, quenching of the redox system with acid; as well as model experimental conditions.[76]

The acceptor Ub monomer was the recombinant Lys to Cys variant of Ub, K48C Ub. The donor Ub was modified at its C terminus with cysteamine. This modification can occur either by incubating UbCOSR with cysteamine at pH 8.0 or by cleaving Ub-intein from a chitin beads column with a 50mM solution of cysteamine at pH 8.0 as well. In both cases the yield remained at 80%. Then, a linear

disulfide was promoted by oxidizing the Ub-cysteamine (UbNHSH) at room temperature and pH 8.0 (Scheme 38).

Donor and acceptor were coupled in the presence of diamide and HEPES as the buffer system, affording the isopeptidic bond mimic of K48Ub$_2$. The ligation is considered finished when the characteristic yellow color of diamide is consumed and by SDS-PAGE since no more monomer is detected. After high performance cation exchange chromatography, three disulfide dimers were isolated; K48C Ub$_2$, (UbNH)$_2$S$_2$, and K48Ub$_2$ S-S, where the latter is the desired dimer (40% yield); as shown in Scheme 39. The profile of elution was also confirmed by ESI-MS as in the case of the thiol-ene chemistry.
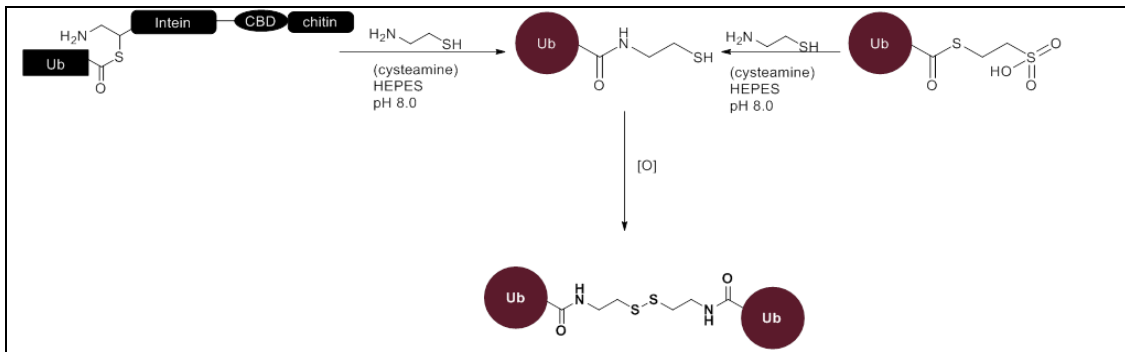
Scheme 38. Ub donor disulfide precursor for intermolecular disulfide exchange.
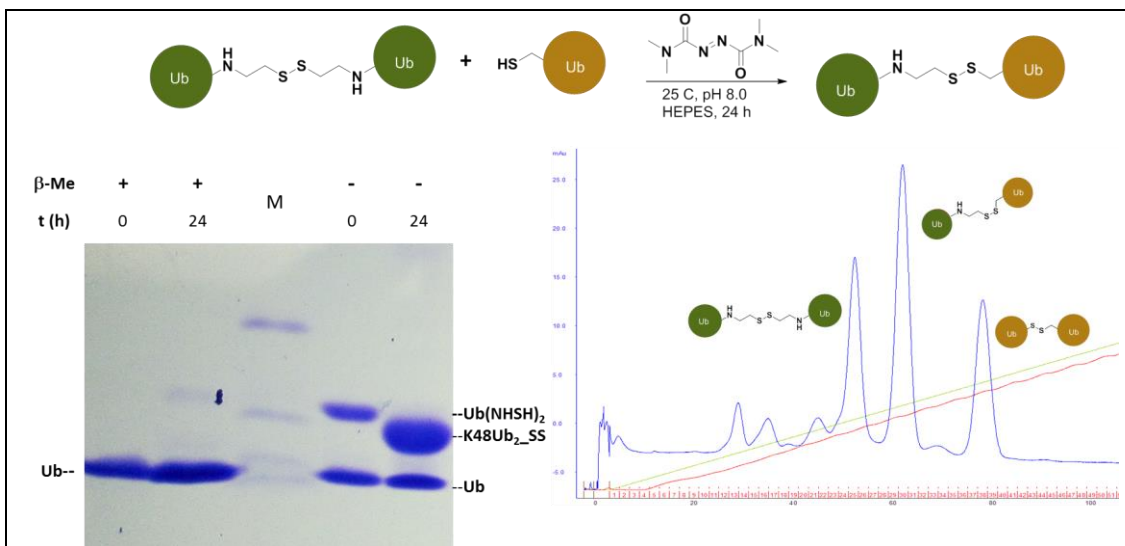
Scheme 39. Intermolecular disulfide exchange in the presence of diamide.

Part C. DUBs assays

The removal of Ub from substrates is carried out by DUBs. The C terminus of Ub as well as some UBLs terminate in a Gly-Gly motif which is recognized by long-known Ub C-terminal hydrolases (UCHs) and Ub-specific processing proteases (UBPs/USPs) and cysteine proteases containing an OTU domain.[78]

One enzyme of each class was used to address the cleavage of these K48Ub$_2$. They were the K48 linkage specific: OTUB1, a universal DUB: USP5 (isoT), and the C-terminal hydrolase: YUH1. One of the immediate questions is if the USP and the OTU cysteine proteases would be as effective as they are with a regular isopeptidic bond, because these isopeptide bond mimics have one more carbon atom and at least a sulfur atom. From the SDS-PAGE, it can be concluded the OTUB1 enzyme was able to start cleaving the dimer after one hour of incubation (Figure 14A, left). The cleavage is almost complete when the incubation time is sixteen hours for the dimer formed by thiol-ene chemistry. The enzyme behaves slightly the same for the dimer formed by disulfide-exchange; though after long exposure no dimer band was detected (Figure 14A, right). However, when USP5 was used, the dimers remained without changes over the course of time (Figure 14B). These results suggest that it is possible the dimers might not acquire the conformation required for the cysteine protease triad to cleave the amide bond; and that lack of activity from the enzyme is due to the fact that the sulfur atom(s) might create a higher steric hindrance or even altering the hydrogen-bonding since their pair of electrons are suitable as Lewis bases. As a positive control, the natural dimer was exposed to the enzymes and the

cleavage occurred almost immediately when incubated with USP5 or OTUB1 (Figure 14C). Finally, as expected YUH1 did not have any effect on the dimers (Figure 14D). This result is consistent with the fact that the active site cysteine in the UCH enzymes lies at the bottom of a narrow groove in the surface of the protein, and it has been reported that larger side chains could not be accommodated in this groove, providing a strong constraint on the substrate selectivity of UCHs. In addition, the substrate moiety must pass through a loop with a very small diameter (no greater than 15 Å), explaining why most UCHs cannot act on Ub–protein conjugates.[79]
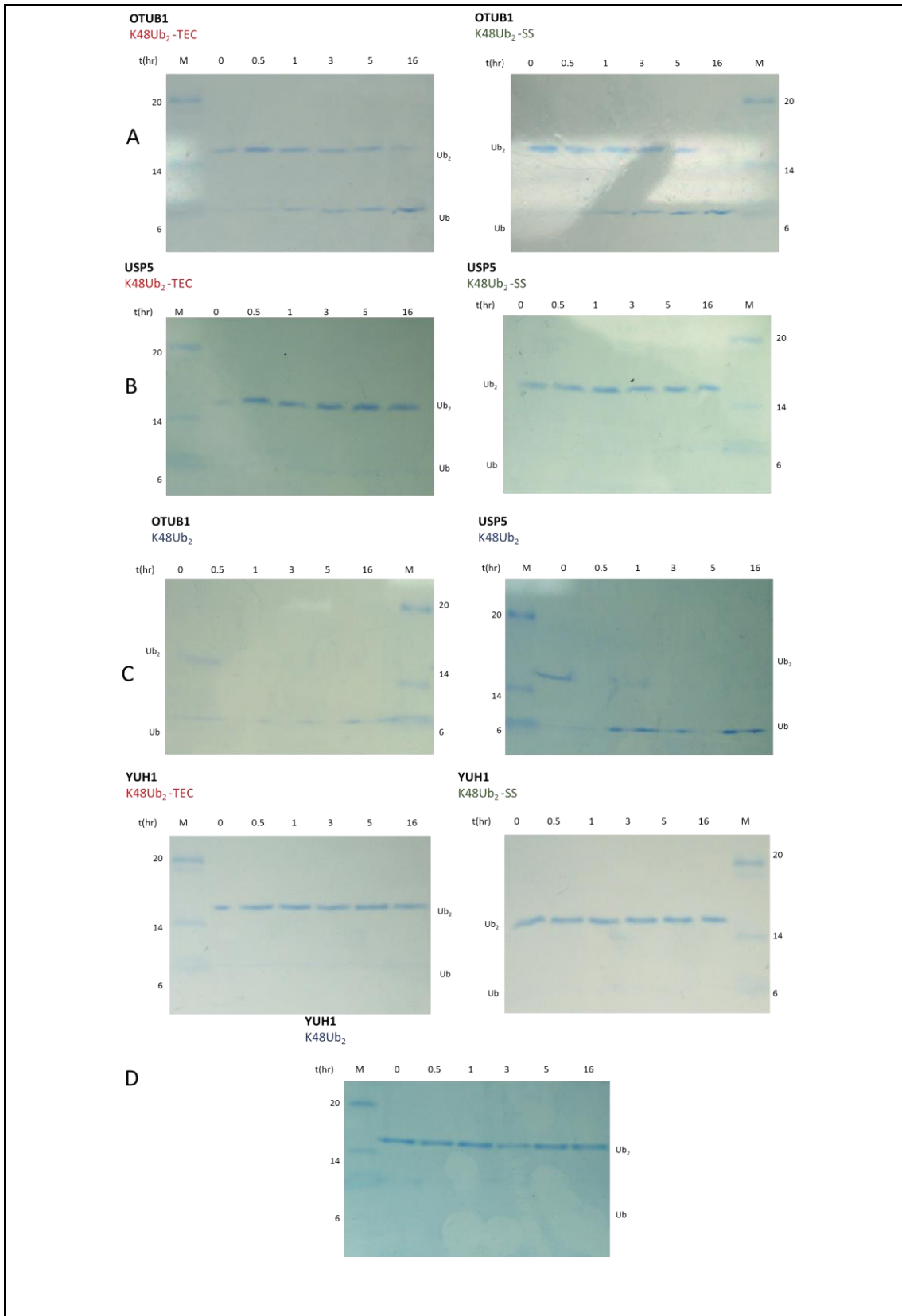
Figure 14. DUB assay of K48Ub₂. A: Incubation with OTUB1; B: Incubation with USP5; C: Positive control with K48Ub2 natural dimer; D: Incubation with YUH1.

Part D. NMR characterization of K48-linked diubiquitin isopeptide bond

mimics

As a second approach the individual domains were first characterized by NMR using a short-time ^{1}H-^{15}N HSQC spectra (referred as SOFAST pulse sequence)[80] at constant temperature (298.1 K) in 20mM phosphate buffer at pH 6.8. Each domain was compared with Ub^{WT} as reference (Figures 15, 16, 17). Additionally when the domains are compared between them, the chemical shift perturbations are around the hydrophobic patch, which is the region between residues L8, I44 and V70. By looking at the signal shifts of residues G10-I13, A46-L50, L67-V70 and the tail L73-G76 there is a tendency towards upfield positions. Considering the signals of the distal domain in the enzymatic K48Ub$_2$ as a starting point reference and the signals of Ub^{WT} as the end-point, the thiol-ene isopeptide mimic shows double signals suggesting the existence of two conformers, one resembling the natural dimer and one suggesting an open conformation. In comparison, the disulfide isopeptide mimic does not exhibit double signals of the hydrophobic patch residues and in addition most of the signals have the propensity to move toward the distal enzymatic chemical shifts suggesting that some degree of interface is present and definitely some flexibility is developed when the linker is extended by one more bond. Therefore, a series of ^{1}H-^{15}N HSQC spectra were recorded at two more different pH values to establish if there is a conformational effect with these isopeptide mimics as it is known for the enzymatic K48-diubiquitin.
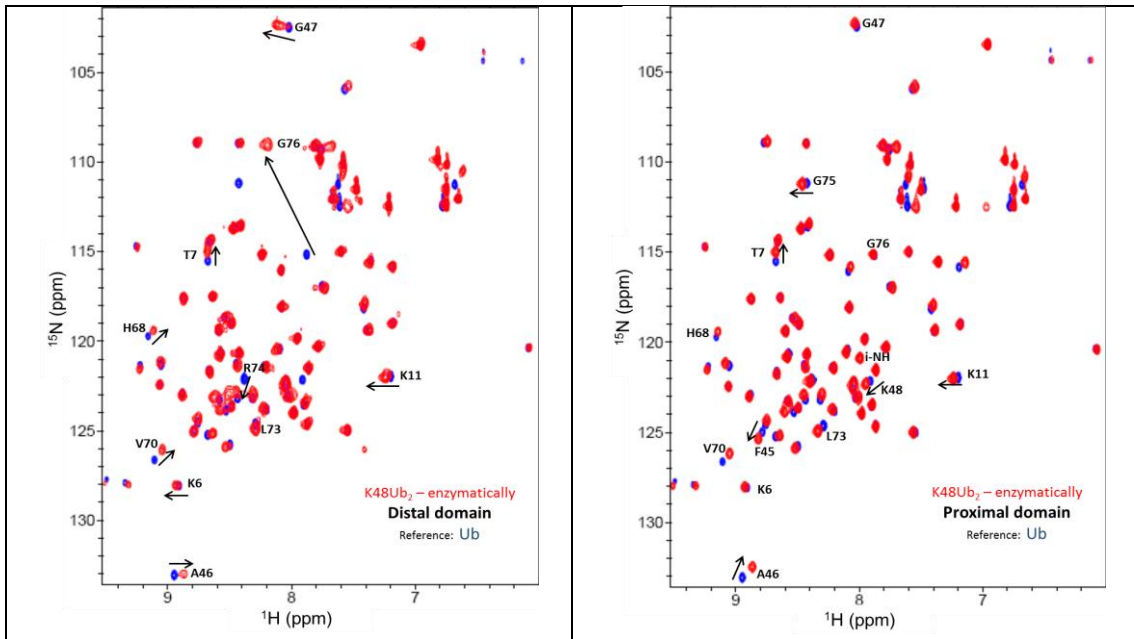
Figure 15. ^{1}H-^{15}N HSQC NMR spectra of K48Ub$_2$ natural dimer. Left: Distal domain; Right: Proximal domain. Blue: Ub, Red: ^{15}N enriched domain.
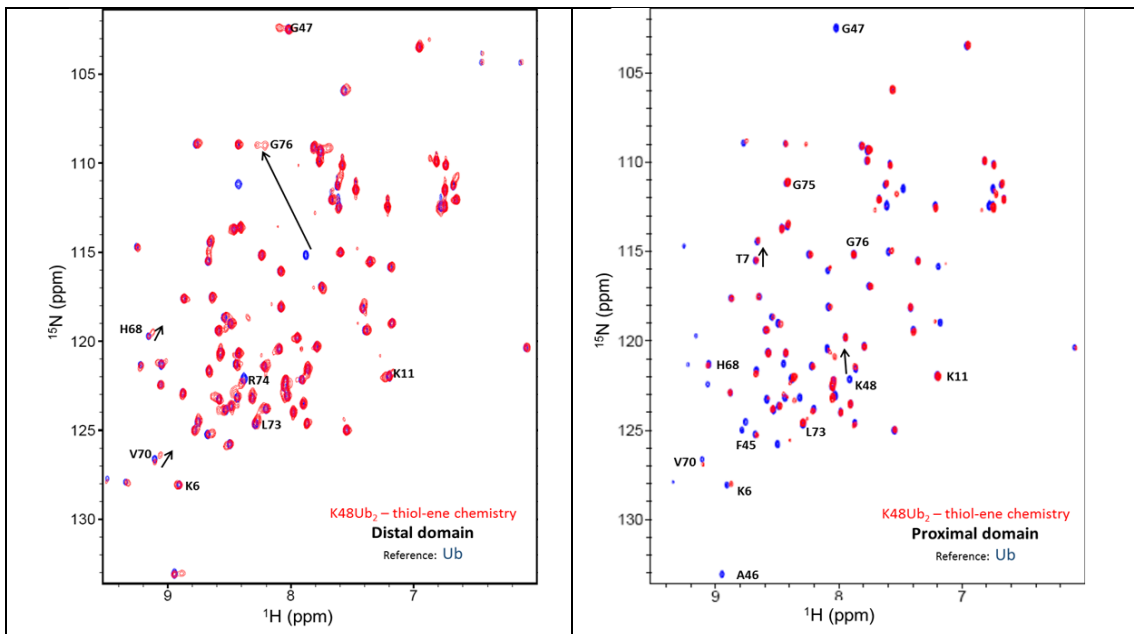
Figure 16. ^{1}H-^{15}N HSQC NMR spectra of K48Ub$_2$ assembled using thiol-ene chemistry. Top: Distal domain; Bottom: Proximal domain. Blue: Ub, Red: ^{15}N enriched domain
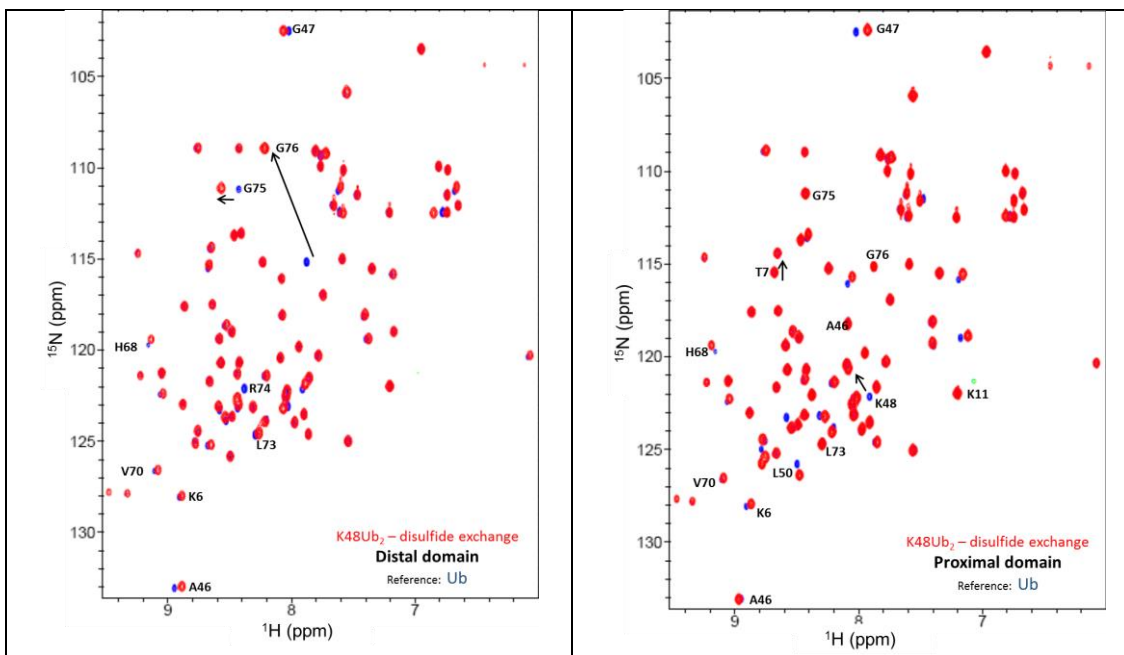
Figure 17. ^{1}H-^{15}N HSQC NMR spectra of K48Ub$_2$ linked via disulfide. Top: Distal domain; Bottom: Proximal domain. Blue: Ub, Red: ^{15}N enriched domain

Part E. Ubiquitin dimers pH-dependent conformational switch

From NMR studies it is well known there are two major conformational states depending on pH in K48Ub$_2$.[9, 81] At near physiological conditions a closed conformation between the domains is adopted and an interface exists between hydrophobic patches of the units. Also, it is known that at acidic pH this interface between the hydrophobic patches is lost.[81, 82] In fact, in a quantitative manner it has been established the population of the open and closed conformations are 80% depending the pH, respectively (Figure 18).[81]

A very useful measurement to keep track of changes in the electronic environment of the nuclei under observation by NMR is following the chemical shift perturbations (CSP). Those residues with significant CSPs are typically involved, for instance, in the interdomain interface in the Ub-Ub linkage. The CSPs are calculated by the

following equation: $\Delta\delta = \left[(\Delta\delta_H)^2 + \left(\Delta\delta_N/5\right)^2\right]^{1/2}$,where $\Delta\delta_H$ and $\Delta\delta_N$ are the differences in chemical shifts of ^{1}H and ^{15}N, respectively, between the dimer and mono Ub.[55] As a control, the enzymatically K48Ub$_2$-linked system was measured first at three different pH values, at pH 4.5 (for the open conformation); at pH 6.8 (where 80% of the population is at closed conformation) and at pH 7.4 (physiological pH); as shown in Figure 18; the open and closed conformations are shown in the ribbon representation.[8, 82] Second, the same measurements were done with the isopeptide mimics and their CSPs were calculated compared to Ub; as shown in Figures 19 and 20, respectively.
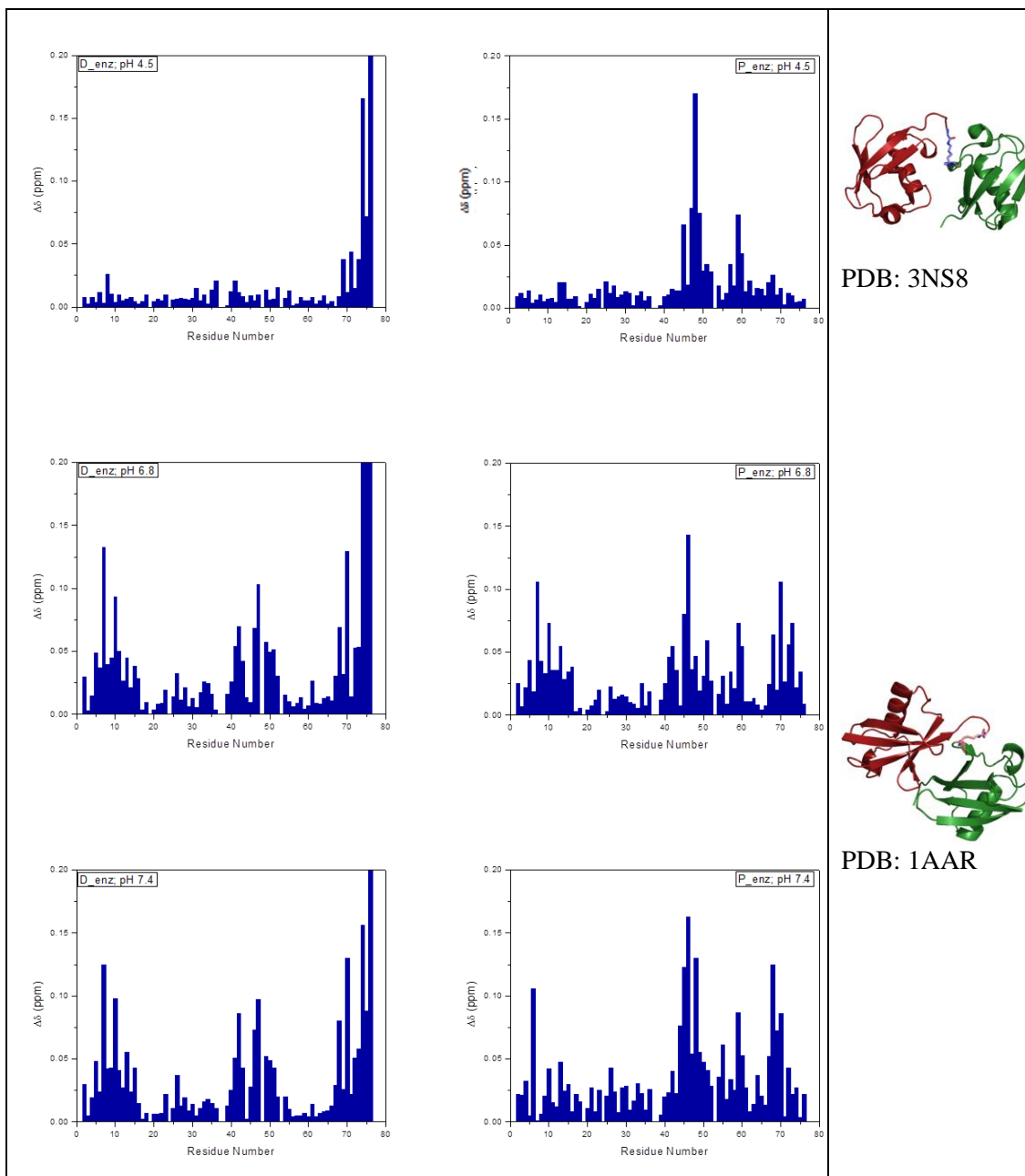
Figure 18. CSPs in K48Ub₂ natural dimer. Left: distal domain; right: proximal domain. The ribbon model of the crystal structures of the open and closed conformation, 3NS8 and 1AAR, respectively.
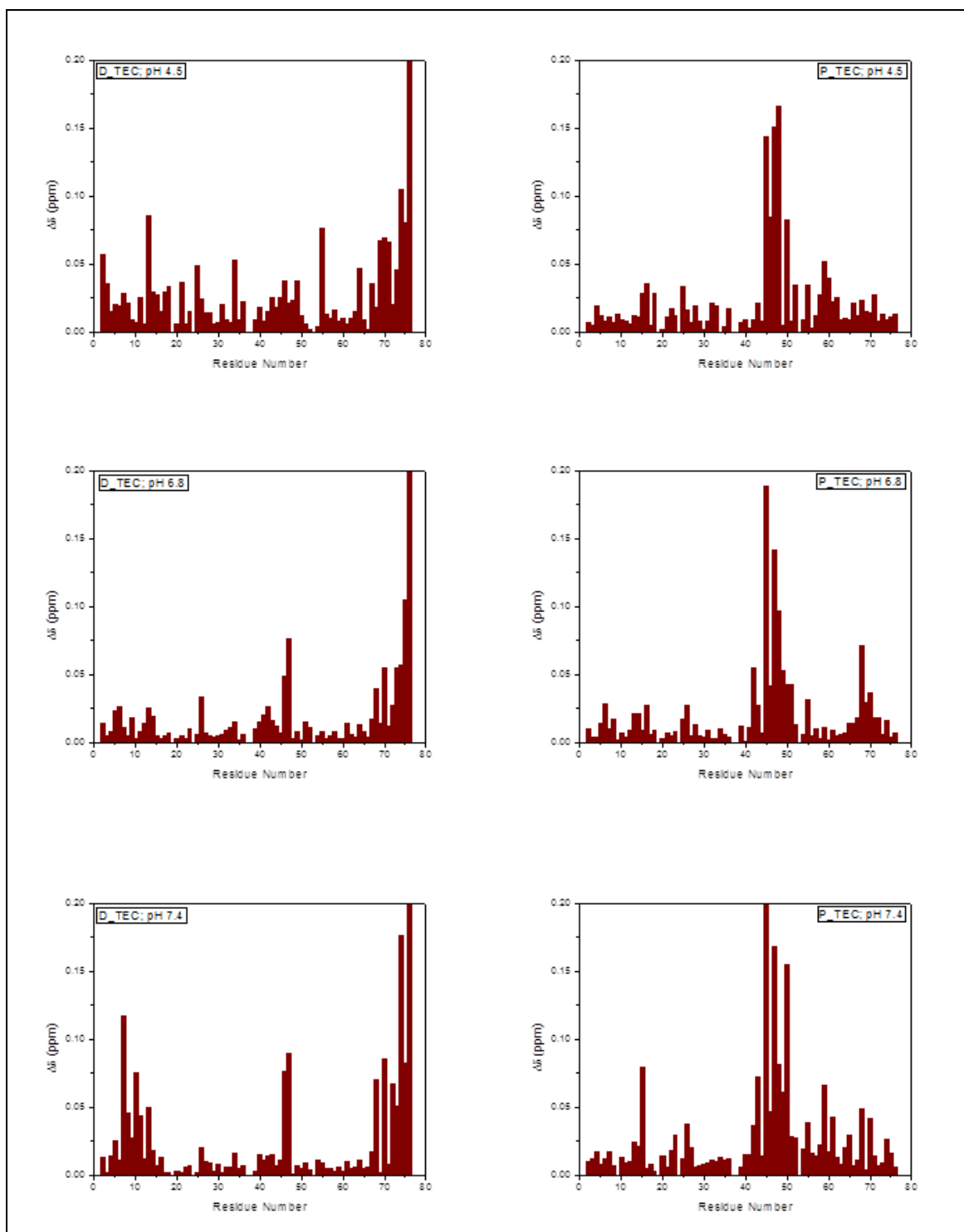
Figure 19. CSPs in K48Ub$_2$ assembled using thiol-ene chemistry; left: distal domain; right: proximal domain.
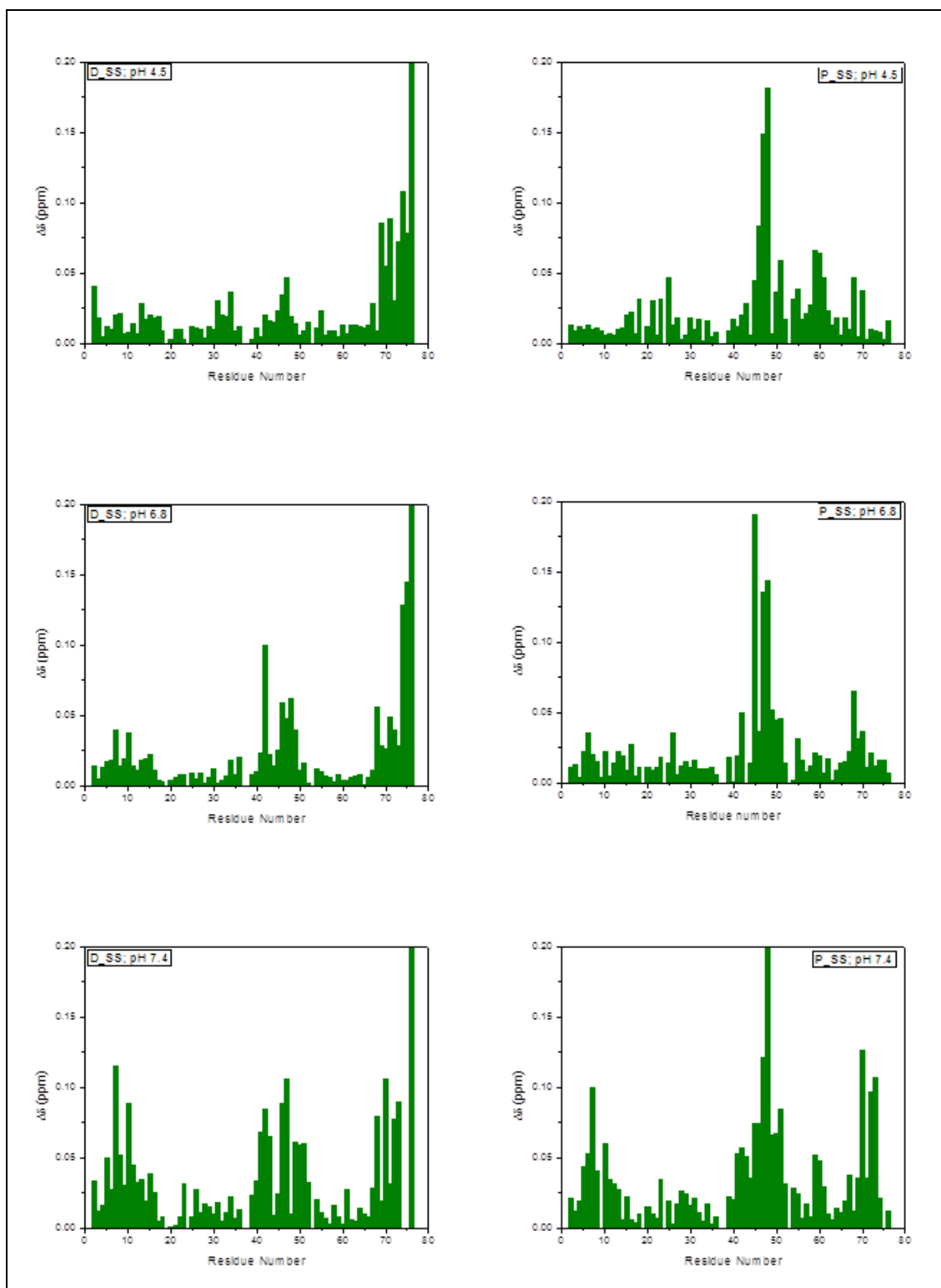
Figure 20. CSPs in K48Ub$_2$ assembled using disulfide-exchange; left: distal domain; right: proximal domain.

The major changes are observed in the distal domain of the dimeric systems. It is noticeable that for the K48Ub$_2$ linked through the thiol-ene chemistry the chemical shift perturbations are located in the region of I13. The interface is more pronounced at acidic pH, and it is still present at pH 7.4, whereas at pH 6.8 the effect could be interpreted as a transition in between the two possible conformations of the dimer. In comparison the proximal domain remained with the chemical shift perturbations located only in the proximity of position 48. The K48Ub$_2$ linked through a disulfide bond shows a different pattern because the interface is clearly present especially at pH 7.4. In addition, the pH dependence is quite comparable to the one observed in the natural dimer; therefore, the S-S bond is less rigid than the C-S bond bringing both domains closer.

In conclusion, two isopeptide bond mimics were synthesized between recombinant Ubs using two water-soluble azo-compounds, V-50 and diamide. Thiol-ene chemistry ligation of a cysteine and allyl systems can be achieved by thermal free radical chemistry, optimizing the reaction conditions when a photochemical initiator or reactor is not available. Disulfide exchange was achieved by redox chemistry taking advantage of the easy exchange diamide provides when the reactive species are at pH 8.0 in HEPES buffer. Moreover, the purification of the dimers by high performance cation exchange chromatography renders the dimers with high purity. NMR characterization of the dimers revealed their similarities and differences from the natural enzymatically synthesized K48Ub$_2$; and it also revealed the 3D structures of each Ub unit are intact. On the one hand, the thiol-ene chemistry dimers showed an interface between Ubs when the pH switches from acidic to slightly alkaline, keeping

89

the perturbations in close proximity to the isopeptide bond suggesting two conformers or a slow transition from an open conformation to a closed conformation. And on the other hand, the disulfide dimers have a similarity with the natural dimers by showing perturbations around the hydrophobic patch and around the isopeptidic bond. Finally, both dimers exhibit different conformational properties, especially the ability to form a closed state due to the pH-dependence of the Ub_2 conformation.

A broader application of the thiol-ene chemistry was carried out with H2A and H2B since these proteins are in an unfolded state when they are not in a complex. However, the proteins are not stable at the pH and temperature which is needed for the thyil radical to make the addition to the allyl system. Nevertheless, the disulfide exchange was tried only once with E2-25K and G76C Ub in the presence of diamide with the purpose of studying the mechanism on how the enzyme can translocate Ub units for dimerization. Because the yield of the reaction was very low, due to the preference of E2-25K to dimerize with itself, the reaction conditions are pending for optimization; but definitely the semisynthetic methodologies are a plausible approach for the covalent ligation of proteins when the enzymatic machinery is not known.

Chapter 4: UBDs and Future Scope

UBDs affinities with Ub~(δSH)K77

  The presence of an extra amino acid residue in Ub by NCL not only worked for the purpose of having a sulfhydryl into the protein which could work for the installation of labels to study the possible conformation changes in the protein. In fact it also changes the environment of the protein in terms of adding a charge since the residue *per se* is a lysine with not only the corresponding free carboxylic acid but also the α-amino moiety free. This opens the question if the binding of a UBD previously characterized with Ub will behave differently when the mercaptolysine is present in the C terminus of Ub. Therefore, a NMR titration experiment with three particular UBDs was carried out. These UBDS, conventionally called ligands, were first the UBA2 from Rad23 (human, HHR23A), a polyUb chain receptor involved in proteasomal degradation modulation; second, the UBA domain of Ubiquilin-1 also implicated in the regulation of the proteasome degradation of cellular proteins. In addition, Ubiquilin-1 is important in the protein quality control system for Huntington's disease;[83] and third the UIM of Rpn10, because its hydrophobic residues are known to recognize the hydrophobic patch of Ub.[9] Rpn10 is involved in the recognition of both proteasome and polyUb chains and acts as a shuttle protein for the ubiquitinated proteins to their degradation fate.[84] In addition, these UBDs have been characterized in our laboratory and their K_D constant with respect to Ub are 4000-500 μM for the UBA2 of Rad23,[85] 20 μM for the UBA of Ubiquilin-1;[86] and ~ 44 μM, for the UIM domain of Rpn10.[87]

For the UBA2 of HHR23A preliminary titration results showed a lower value of the K_D. Going from 5:1 [L/P] ratio, a Kd ~ 200 μM was calculated when the UBA was titrated into ^{15}N~Ub(δSH)K77. In comparison, when the UbaY of Ubiquilin-1 was used the K_D ~ 18 μM. These results suggest the binding is tighter for UBA2 while the binding for UbaY remained in the same range as the previous reported for Ub. In terms on how the ligand is interacting with Ub it can be proposed that these ligands are interacting as expected with the hydrophobic patch of Ub. The CSPs at the end point of titration are consistent with those previously reported for Ub, though it is interesting to notice that N25 was a particular residue that had the strongest perturbation. Finally, the K_D for the UIM of Rpn10 as a preliminary titration result had a value ~20 μM. Again a diminishing by half its value with respect of Ub, though not a decrease in order of magnitude, which is consistent with the idea that the binding is tighter between the hydrophobic patches and the extra charge caused by the presence of the mercaptolysine may help in a better conformation for the interfaces to interact. In addition, the model for fitting the K_D suffices a 1:1 ligand-protein stoichiometry for all three UBDs.
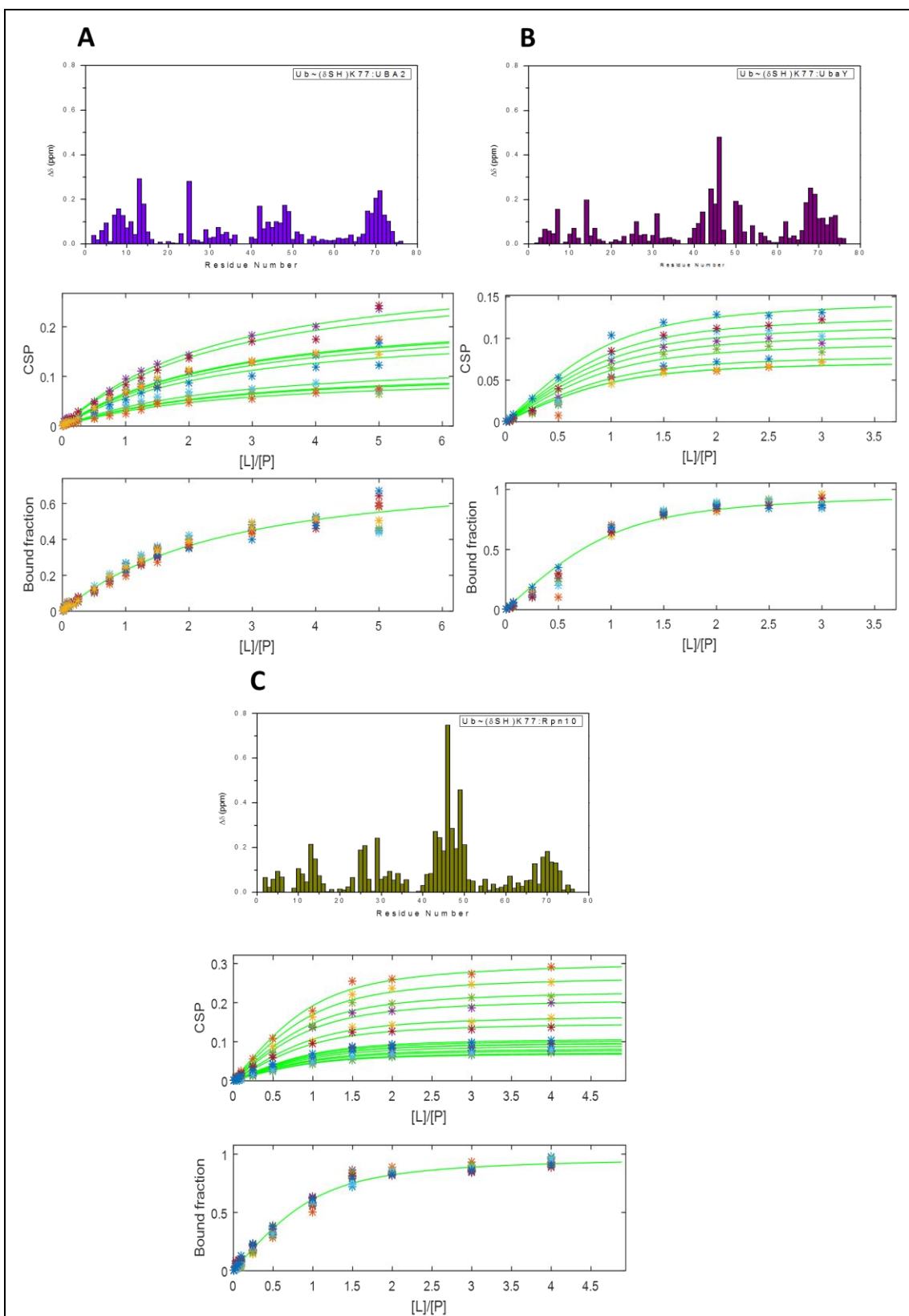
Figure 21. Characterizing the Ub~(δSH)K77 interactions with UBDs. A: UBA2 (HHR23A), B: UbaY (Ubiquilin-1), C: UIM (Rpn10). Plot bars are CSPs at the final point of titration, curves show the output of in-house program for K_D determination (KDfit).

93

Trifunctional probes for inhibitor screening as possible pharmacological

therapeutics

From the point view of therapeutics, research has focused the energy in developing unsaturated C-H, C-O systems in conjugation with Ub for scavenging DUBs or the E2-E3 enzymes in the UPS (from Michael acceptors, cross-linkers to π-stacking molecules) as therapeutic inhibitors.[88] The rationale behind it is if a misfolded protein will aggregate, it will be tagged for degradation by Ub, as the first stage of control; but the degradation pathway won't be efficient since the activity of it is affected by longevity. Therefore, the inhibition of some stages alleviates the aggregation of misfolded proteins, decreasing the cytotoxicity effect.[89] A very interesting model for screening small molecules *in vitro* is based on a trifunctional probe.[90] The first part of the probe is a molecule where the potential inhibitor is part of the scaffold, the second part a fluorescent or affinity tag and the third part is the protein of interest with the reaction motif, for instance Ub (or UBDs) because they are potentially easy to purify, and used for characterize the efficiency of inhibition or binding properties when the probe reacts either with the DUB or the E2-E3 enzymes or other UBDs.

With this idea, a covalent system where the small molecule (or inhibitor) bears an α,β-unsaturated atom with an electrowithdrawing group (EWG) in one edge and a nucleophile (Figure 22A) or a good leaving group (LG) (Figure 22B) can be proposed. Ub~(SH)K77 (monomer or as the proximal domain of a Ub chain, which linkage could be customized, but to keep consistent K48 as a first approach), where

the sulfhydryl group is the reactive functional group, is used as a constant binding motif. And a protein of study which could be a) UBD bearing a good leaving group (Figure 23A) or if it is the case with a nucleophile (Figure 23B) or b) DUB, E2-E3 with the cysteine of their corresponding active site (Figure 23B) which in any case would have either a fluorescent tag or an affinity tag for purification.
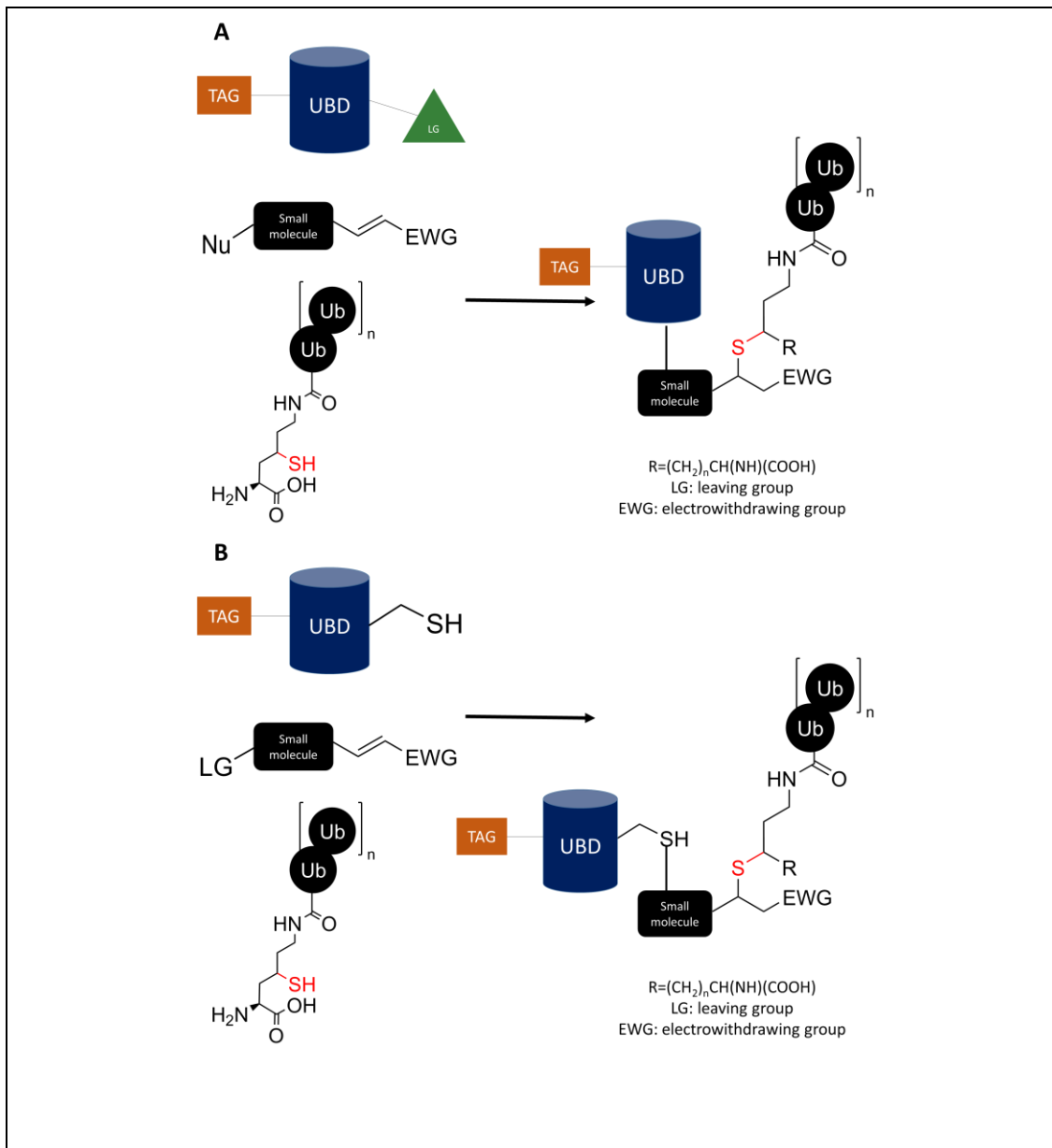
Figure 22. Two scenarios of trifunctional probes. A: The nucleophile is in the small molecule and the leaving group in the protein of study. B: The nucleophile is in the protein of study and the leaving group is in the small molecule. Ub~(SH)K77 is used as a constant binding molecule and it could be polymerized, if needed.

95

Nevertheless, the presence of the mercaptolysine is not a mandatory feature for the trifunctional probe. Ub can be modified in its C terminus by mutation of G76 to G76C, or by reaction with cysteine in solution by NCL to afford UbC77 with the same experimental conditions where the protein is note exposed to denaturating conditions during NCL. In addition, the dimer or polyUb chain (if it is the case) is not limited to a natural dimer, the non-canonical isopeptide bonds could be also used as a point to characterize the recognition or binding properties of UBDs.

Synthesis of K48Ub$_2$ by thiolysine

A controlled enzymatic approach for isopeptide bond mimics has been previously reported where the proximal domain is activated first as a thiolysine and the distal domain was a K48R mutation.[56, 91] As described in the natural dimer section, the distal domain with a Boc-lysine introduced in Ub by site-directed mutagenesis was used;[50] and the proximal domain has an aspartic acid residue which will be cleaved by YUH1 hydrolysis. To activate the proximal domain and mimic a lysine residue, K48C Ub was alkylated with bromoethylamine.[91] In contrast with TEC and disulfide exchange, this isopeptide bond mimic has the same length as the C-C backbone present in lysine but keeping one carbon substituted by a sulfur atom (Scheme 40). After high performance cation exchange chromatography the dimer was isolated. From 5 mg of proximal monomer, approximately 5 mg of dimer were isolated (50% yield) recovering 30% of proximal domain with the G76 residue exposed, which indicates that the thiolysine is not a specific substrate for the E2. In order to study the dimer easily by NMR each domain was ^{15}N isotopically labeled and

ligated with its not isotopically enriched counterpart separately. The CSPs shows a similar behavior with respect the natural dimer at pH 6.8 (Figure 23) for both domains when compared with Ub^{WT}. The conformational pH switch dependence as well as studying how these dimers behaves when titrated with UBDs such as UBA2 (HHR23A), and Rpn10/Dsk2 which has higher affinity for polyUb chains is a question that remains open for further research.
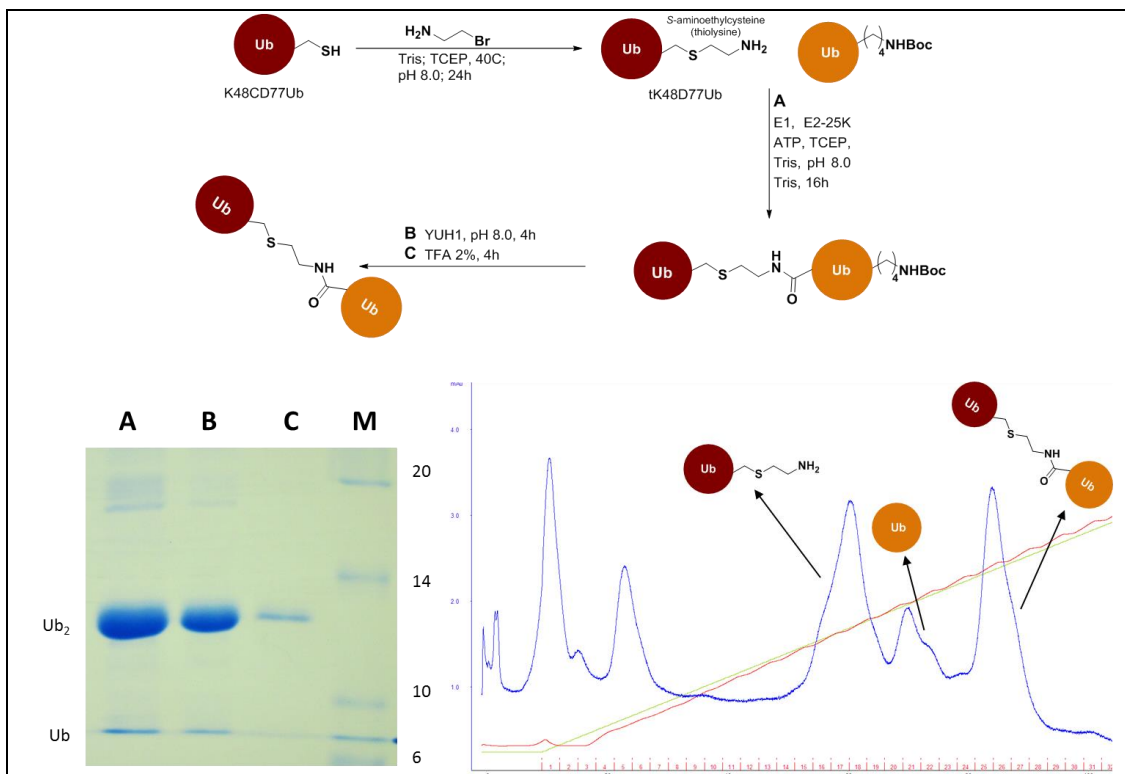
Scheme 40. Synthesis of K48Ub$_2$_thiolysine (tK).
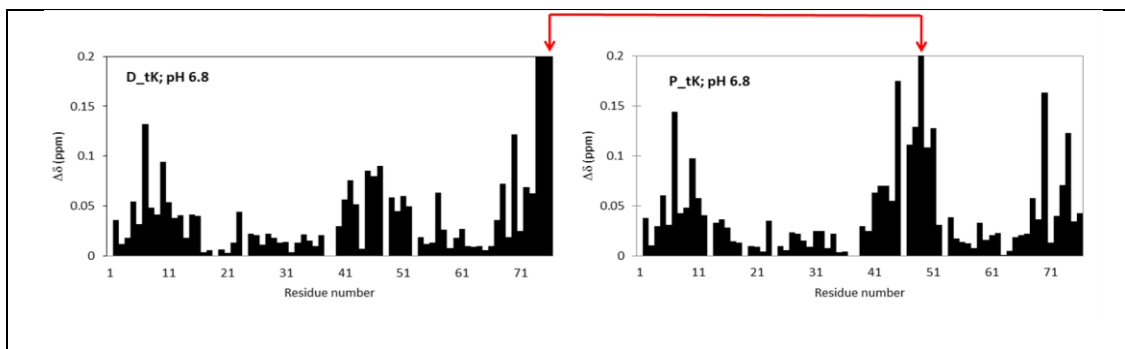
Figure 23. CSPs of K48Ub$_2$, assembled enzymatically with thiolysine at pH 6.8. Left: distal domain, Right: proximal domain. The red arrow indicates which residues form the isopeptide bond.

Chapter 5: Materials and Methods

Section 1. Synthesis of mercaptolysines

General Methods

All chemicals used in this experimental procedure were purchased at Sigma, Acros, or Fisher. The solvents used in the synthesis were distilled and stored over molecular sieves to avoid the presence of water. The solvents used for column chromatography were used without any previous purification step. Precoated silica gel plates with a fluorescent indicator (Merck 60 F_{254}) were used for thin layer chromatography (TLC). Silica gel column chromatography was carried out with Alfa Aesar silica gel (60-230 mesh). ^{1}H NMR spectra of the synthetic route towards the amino acids were recorded at 400 MHz and ^{13}C NMR spectra were recorded at 100 MHz on Bruker AV-400 MHz. The chemical shifts are expressed in ppm from internal solvent peaks (CDCl$_3$, 7.24 ppm; d$_6$-DMSO, 2.50; D$_2$O, 4.85; ^{1}H-NMR and CDCl$_3$, 77.00 ppm; d$_6$-DMSO, 31.25 ppm; ^{13}C-NMR). Coupling constant values are given in Hertz (Hz). The coupling patterns are expressed by s (singlet), d (doublet), t (triplet), m (multiplet), brs (broad singlet), brd (broad doublet), etc. High resolution mass spectra were measured on a JEOL AccuTOF-ESI-MS mass spectrometer in electrospray positive mode using flow injection.

In terms of protein expression, LB media was used for starter cultures. Auto-inducing media was used for protein expression of non-labeled protein and 15N auto-inducing minimal media variation was used for 15N-isotopically labeled protein. All dialysis and buffer exchanges were performed using 3.5kDa MWCO membranes (Thermo

Scientific and Millipore). Protein chromatography was performed in an ÄKTA plus system.

. 1H, 15N NMR spectra of any protein sample were recorded at 800 MHz, 81 MHz on Bruker Ascend-800; and at 600 MHz, 61 MHz on Bruker UltraShield-600.

Synthesis of γ-mercaptolysine
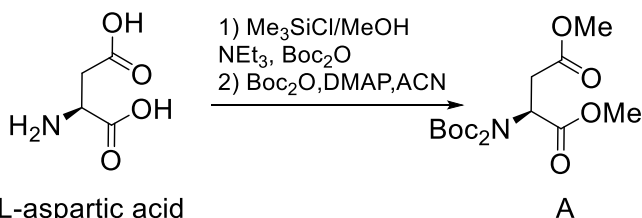
L-aspartic acid   A

Dimethyl (S)-2-*tert*-butxycarbonylamino-butanodioate: A solution of aspartic acid (5.00 g, 37.57 mmol) in 125.00 mL of methanol was cooled down to 0 °C. Then, trimethyl chlorosylane (21.00 mL, 165.29 mmol) was added slowly at this temperature. The reaction was then stirred over two days. Triethylamine (33.54 mL, 244.17 mmol) was added slowly to the reaction mixture, following by the addition of Boc anhydride (9.35 mL, 41.32 mmol) and left stirring over 16 hours at room temperature. After a second addition of Boc anhydride (4.70 mL, 20.35 mmol) and sixteen more hours at room temperature (*Rf: 0.5 Hex:EtOAc 75:25*) the methanol was removed under vacuum from the reaction mixture. Then, the reaction mixture was extracted with ethyl acetate (3 x 40 mL), dried over anhydrous sodium sulfate and the ethyl acetate was removed under vacuum and the product was isolated as an orange oil (95%). ^{1}H NMR (CDCl$_3$, 400 MHz) δ 1.40 (s, 9H), 2.88 (ddd, 4.56, 17.01, 21.69 Hz, 2H), 3.66 (s, 3H), 3.72 (s, 3H), 4.54 (m, 1H) 5.47 (d, 7.73 Hz, 1H). ^{13}C NMR (CDCl$_3$, 100 MHz) δ 27.99, 35.63, 52.59, 54.97, 83.65, 151.52, 171.03, 173.07. ESIMS for C$_{11}$H$_{19}$NO$_6$ [MH$^+$]: 262.12. Found: 262.2983.

Dimethyl (2S)-2-(*tert*-butoxy)-N-((*tert*-butyl)oxycarbonyl)carbonylaminobutane-1,4-dioate (**A**): A solution of Dimethyl (S)-2-*tert*-butxycarbonylamino-butanodioate (8.99

g, 34.41 mmol) was dissolved in acetonitrile (115.00 mL). DMAP (0.84 g, 6.88 mmol) was then added to the reaction mixture, followed by Boc anhydride (8.70 mL, 37.85 mmol). After 2 hours of stirring, a second addition of Boc anhydride (4.35 mL, 18.92 mmol) was done and then left stirring over 16 hours at room temperature. The reaction mixture was extracted with ethyl acetate (3 x 30 mL) and washed with water (20 mL) and brine (20 mL); dried over anhydrous sodium sulfate and the ethyl acetate was removed under vacuum. The orange oil isolated was purified by column chromatography (*Rf: 0.75 Hex:EtOAc 80:20*) and a crystalline oil was isolated (90%). ^{1}H NMR (CDCl3, 400 MHz) δ 1.47 (s, 18H), 2.70 (dd, 6.51, 16.46 Hz, 1H), 3.21 (dd, 7.09, 16.45 Hz, 1H), 3.67 (s, 3H), 3.69 (s, 3H), 5.44 (td, 6.82, 6.82, 13.63 Hz, 1H). ^{13}C NMR (CDCl$_3$, 100 MHz) δ 27.95, 28.28, 35.68, 51.95, 54.88, 83.55, 151.54, 170.29, 171.05. ESIMS for C$_{16}$H$_{27}$NO$_8$ [MH$^+$]: 362.39. Found: 362.4143.
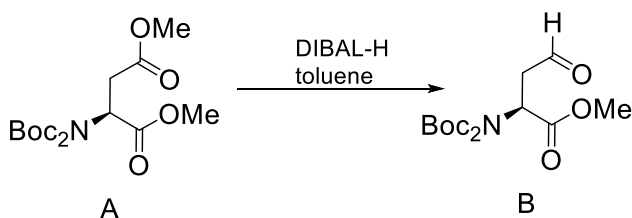
Methyl (2S)-2-(*tert*-butoxy)-N-((*tert*-butyl)oxycarbonyl)carbonylamino-4-oxobutanoate (**B**): A solution of **A** (10.84 g 29.99 mmol) in freshly distilled and dry toluene (150.00 mL), was flushed with argon for 30 minutes. Then, it was cooled down to -78 °C (dry ice - acetone bath). DIBAL-H (19.34 mL of 1.2 M in toluene, 23.19 mmol) was added slowly and dropwise via syringe for over 15 minutes. The reaction was left stirring for 4 hours at -78 °C until no further formation of aldehyde was detected by TLC (*Rf: 0.45 Hex:EtOAc 80:20*). At -78 °C, the reaction was then quenched with 3 volumes of methanol, and left stirring until it warmed up to -20 °C. Then, the reaction mixture was stirred vigorously with 5 volumes of a saturated solution of Rochelle's salt for over an hour. Finally, it was extracted with ether (3 x 70 mL). The ethereal phase was washed with brine (1 x 50 mL), dried over anhydrous sodium sulfate and the ether was removed under vacuum. A yellow oil was purified by column chromatography (*Rf: 0.4; 80:20 Hexane/Ethyl Acetate*). Aldehyde **B** was isolated as a colorless oil (80%). ^{1}H NMR (CDCl$_3$, 400 MHz) δ 1.46 (s, 18H), 2.81

(dd, 5.93, 17.92 Hz, 1H), 3.31 (ddd, 7.00, 17.19, 23.65 Hz, 1H), 3.70 (s, 3H), 5.47 (m, 1H) 9.77 (t, 1.16 Hz, 1H). ^{13}C NMR (CDCl$_3$, 100 MHz) δ 27.94, 44.96, 52.63, 52.85, 83.72, 151.62, 170.25, 198.45. ESIMS for C$_{15}$H$_{25}$NO$_7$ [MH$^+$]: 332.36. Found: 332.3582.
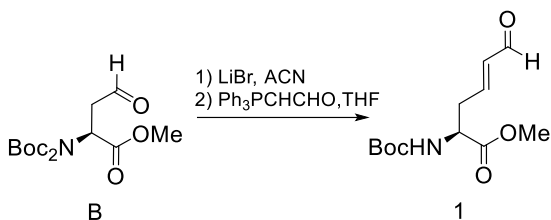
Methyl (2S)-((*tert*-butoxycarbonyl)amino)-4-oxobutanoate: Aldehyde **B** (7.60 g, 22.94 mmol) was dissolved in acetonitrile (261.00 mL). LiBr (5.98 g, 68.82 mmol) was added, and the reaction mixture was heated at 65 °C, over 16 hours. The acetonitrile was removed under vacuum and a light yellow oil was isolated (90%), and stored at 4 °C. (*Rf: 0.15, Hex:EtOAc 80:20*). ^{1}H NMR (CDCl$_3$, 400 MHz) δ 1.46 (s, 9H), 3.09 (m, 2H), 3.77 (s, 3H), 4.64 (m, 1H), 5.44 (brd, 7.40 Hz, 1H) 9.76 (s, 1H). ^{13}C NMR (CDCl$_3$, 100 MHz) δ 28.21, 45.99, 52.75, 80.30, 155.29, 171.46, 199.41. ESIMS for C$_{10}$H$_{17}$NO$_5$ [MH$^+$]: 232.25. Found: 232.2146.

Methyl (2S)-N-(*tert*-butoxycarbonyl)-2-amino-6-oxohex-4-enoate (**1**): Aldehyde **B** (4.42 g, 19.08 mmol) was dissolved in 381 mL of dry THF. To the solution, triphenylphosphoranylidene-acetaldehyde (6.39 g, 20.98 mmol) was added. The reaction mixture was heated at 35 °C for 48 hours under argon atmosphere. At room temperature, the THF was removed under vacuum, and extracted with ethyl acetate (3 x 30 mL). The organic phase was washed with brine (30 mL) and dried over anhydrous sodium sulfate. A red oil was adsorbed in 5 g of silica. Purified by column chromatography (*Rf: 0.4 Hex:EtOAc 70:30*) with a final yield of 70%. ^{1}H NMR (CDCl$_3$, 400 MHz) δ 1.41 (s, 9H), 2.48 (m, 1H), 2.70 (m, 1H), 2.89 (m, 1H), 3.75 (s, 3H), 4.53 (q, 6.77 Hz, 1H), 5.12 (brd, 7.67 Hz, 1H), 6.15 (m, 1H), 6.73 (m, 1H), 9.48 (d, 7.83 Hz, 1H). ^{13}C NMR (CDCl$_3$, 100 MHz) δ 28.24, 30.30, 36.04, 52.73, 68.53, 80.45, 125.50, 135.59, 151.32, 171.59, 193.36. ESIMS for C$_{12}$H$_{19}$NO$_5$ [MH$^+$]: 258.28. Found: 258.2813.
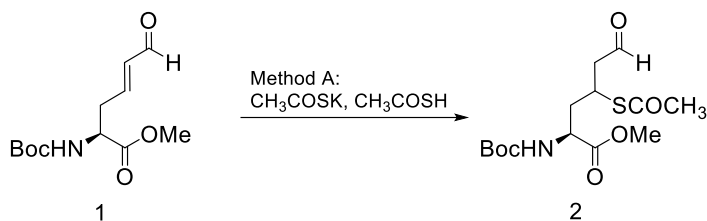
Method A: (2S)-Methyl 4-acetylsulfanyl-2-((*tert*-butoxycarbonyl)amino)-6-oxo-hexanoate (**2**): Aldehyde **1** (3.09 g, 12.02 mmol) was dissolved in thioacetic acid (20.00 mL). This solution was kept at 10 °C. In a separate container under argon and at 10 °C; potassium thioacetate (2.75 g , 24.04 mmol) and 8.00 mL of thioacetic acid were mixed and left stirring for 10 minutes. To the thioacetate solution, the aldehyde solution was added via syringe. The reaction was left stirring for three hours. Then, the reaction mixture was left to reach room temperature and diluted with ether (30 mL). Then, the reaction mixture was extracted with 3 volumes of water (90 mL), the ethereal phase was dried over anhydrous sodium sulfate and the ether removed under vacuum. A dark red oil was isolated (70%) after column chromatography (*Rf: 0.40 Hex:EtOAc 75:25*). ^{1}H NMR (CDCl$_3$, 400 MHz) δ 1.48 (s, 9H), 2.30 s (3H), 2.61 (m, 2H), 2.85 (m, 2H), 3.08 (m, 1H), 3.76 (s, 3H), 4.80 (m, 1H), 5.08 (brdd, 7.18, 14.36 Hz, 1H), 9.70 (s, 1H). ^{13}C NMR (CDCl$_3$, 100 MHz) δ 26.46, 26.81, 27.84, 29.99, 36.29, 52.34, 52.54, 64.09, 82.23, 84.05, 153.47, 171.41, 193.46, 196.25. ESIMS for C$_{14}$H$_{23}$NO$_6$S [MH$^+$]: 334.40. Found: 334.4213.
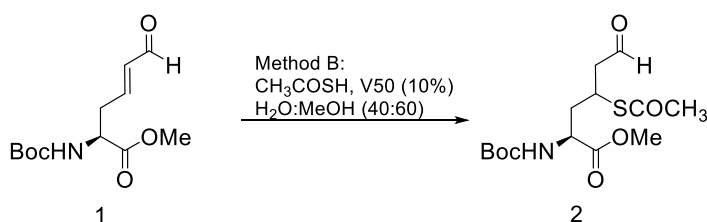
Method B: (2S)-Methyl 4-acetylsulfanyl-2-((*tert*-butoxycarbonyl)amino)-6-oxo-hexanoate (**2**): Aldehyde **1** (3.00 g, 11.66 mmol) was dissolved in methanol (116.00 mL) at room temperature. Then thioacetic acid was added (1.64 mL, 23.32 mmol). In a separate container at room temperature V-50 initiator (0.32 g, 1.17 mmol) was dissolved in water (78.00 mL). The aqueous solution of V-50 was then added to the methanolic solution via syringe and heated up at 50 °C for a period of ten hours. Then, the reaction mixture was left to reach room temperature and the solvent was

removed under vacuum. Redissolved in ether (20 mL) it was extracted with 5 volumes of water (100 mL), the ethereal phase was dried over anhydrous sodium sulfate and removed under vacuum. A dark red oil was isolated (70%) after column chromatography (*Rf: 0.40 Hex:EtOAc 75:25*). ^{1}H NMR (CDCl$_3$, 400 MHz) δ 1.48 (s, 9H), 2.30 s (3H), 2.61 (m, 2H), 2.85 (m, 2H), 3.08 (m, 1H), 3.76 (s, 3H), 4.80 (m, 1H), 5.08 (brdd, 7.18, 14.36 Hz, 1H), 9.70 (s, 1H). ^{13}C NMR (CDCl$_3$, 100 MHz) δ 26.46, 26.81, 27.84, 29.99, 36.29, 52.34, 52.54, 64.09, 82.23, 84.05, 153.47, 171.41, 193.46, 196.25. ESIMS for C$_{14}$H$_{23}$NO$_6$S [MH$^+$]: 334.40. Found: 334.4123.
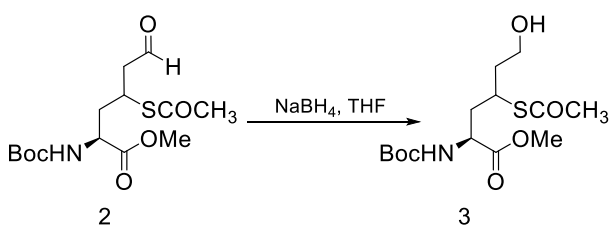
(2S)-Methyl 4-acetylsulfanyl-2-((tert-butoxycarbonyl)amino)-6-hydroxy-hexanoate (**3**): Aldehyde **2** (2.45 g, 7.36 mmol) was dissolved in 150.00 mL of anhydrous THF. The system was cooled down to -10 °C. NaBH$_4$ (0.56 g, 14.80 mmol) was added to the reaction very slowly over a period of 20 minutes and monitored by TLC. After no further advance of the reaction, it was quenched at -10 °C with 10 mL of NH$_4$Cl saturated. Extracted with ethyl acetate (3 x 20 mL) and dried over anhydrous sodium sulfate. The ethyl acetate was removed under vacuum and the product was purified by column chromatography (*Rf: 0.4 Hex:EtOAc 65:35*). A clearless oil was isolated (70%) and characterized. ^{1}H NMR (CDCl$_3$, 400 MHz) δ 1.43 (s, 9H), 1.57 (m, 2H), 1.84 (m, 1H), 2.31 (s, 3H), 3.13 (m, 1H), 3.38 (m, 1H), 3.66 (m, 1H), 3.78 (s, 3H), 4.80 (m, 1H), 5.40 (dd, 7.7, 57.2 Hz, 1H). ^{13}C NMR (CDCl$_3$, 100 MHz) δ 27.76, 28.22, 38.27, 38.76, 41.28, 51.32, 51.90, 58.82, 81.48, 155.30, 171.75, 195.28. ESIMS for C$_{14}$H$_{25}$NO$_6$S [MH$^+$]: 336.42. Found: 336.4223.
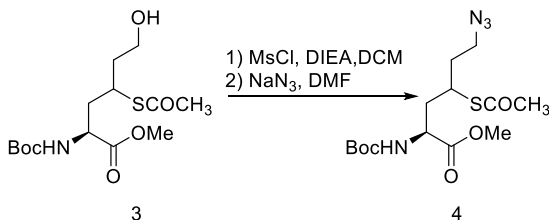
(2S)-Methyl 4-acetylsulfanyl-6-azido-2-((*tert*-butoxycarbonyl)amino)-hexanoate (**4**): Alcohol **3** (1.38 g, 4.11 mmol) was dissolved in 5.00 mL of DCM. DIEA (2.14 mL, 12.33 mmol). Then the reaction mixture was cooled down to 0 °C and MsCl (0.48 mL, 6.16 mmol) was added dropwise to the reaction. The solution was left stirring for 6 hours and then extracted with DCM (3 x 20 mL). The organic phase was washed with water (10 mL), brine (10 mL), and dried over anhydrous sodium sulfate. Finally, the DCM was removed under vacuum and the product (90%, *Rf: 0.60 Hex:EtOAc 65:35*) was isolated as a clear yellow oil which was used in the next step without further purification. ESIMS for $C_{15}H_{27}NO_8S$ [MH$^+$]: 414.51. Found: 414.5128. The mesylate was dissolved in 20.00 mL of anhydrous DMF. To the solution, NaN$_3$ (0.48 g, 7.40 mmol) was added and the system was heated up at 80 °C for 8 hours. Then diluted with 5 volumes of water and extracted with ethyl acetate (3 x 10 mL). The ethyl acetate was dried over anhydrous sodium sulfate and removed under vacuum. The product was then purified by column chromatography (85%, *Rf: 0.70 Hex:EtOAc 65:35*). ^{1}H NMR (CDCl$_3$, 400 MHz) δ 1.43 (s, 9H), 1.70 (dd, 6.8, 13.1 Hz, 2H), 1.90 (m, 2H), 2.31 (s, 3H), 3.15 (m, 2H), 3.78 (s, 3H), 3.95 (m, 1H), 4.14 (dd, 7.8, 13.3 Hz, 1H), 5.30 (d, 7.6 Hz, 1H). ^{13}C NMR (CDCl$_3$, 100 MHz) δ 27.80, 28.25, 35.15, 36.44, 38.48, 47.43, 51.79, 68.93, 81.66, 155.24, 171.62, 196.28. ESIMS for $C_{14}H_{24}N_4O_5S$ [MH+]: 361.43. Found: 361.4323.
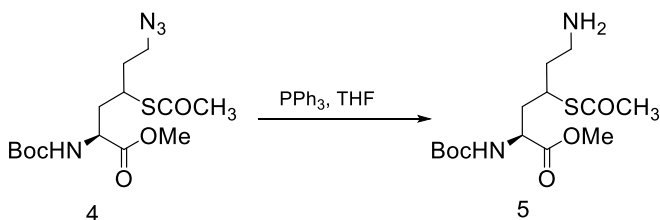
(2S)-Methyl 4-acetylsulfanyl-6-amino-2-((*tert*-butoxycarbonyl)amino)-hexanoate (**5**): Azide **4** (0.78 g, 2.17 mmol) was dissolved in 45 mL of anhydrous THF. At room temperature, triphenylphosphine (0.85 g, 3.25 mmol) was added. Then, the system was heated under reflux for 3 hours. At room temperature, 190 mL of water were added and left stirring one more hour. The reaction mixture was then extracted with ether (3 x 20 mL) and washed with brine (10 mL). The organic phase was dried over anydrous sodium sulfate and the ether was removed under vacuum. The amine (*Rf:*

0.15 Hex:EtOAc 70:30) was stored at 4°C as a yellow oil. ESIMS for $C_{14}H_{26}N_2O_5S$ [MH$^+$]: 335.43. Found 335.4435.
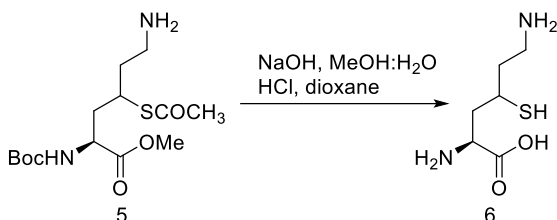
(2S)-2,6-diamino-4-mercapto hexanoic acid (**6**): Amine **5** (0.73 g, 2.17 mmol) was dissolved in 18.55 mL of MeOH:H_2O (2:1). To the solution was added NaOH 1N (9.76 mL, 9.76 mmol) via syringe over a period of ten minutes and under argon atmosphere. After five hours of reaction, the system was cooled down to -10 °C, and HCl 12N in dioxane was added (0.90 mL, 10.85 mmol) under argon. The system was left stirring at this temperature for six hours and then only one hour at room temperature. The solvent was then removed under vacuum and the amino acid was triturated with ether (15mL), and stored at 4 °C. ^{1}H NMR (D_2O, 400 MHz) δ 1.97, 2.09 (m, 4H), 2.89 (m, 1H), 3.30 (t, 6.39 Hz, 2H) 3.83 (s. 1H). ^{13}C NMR (D_2O, 100 MHz) δ 30.36, 31.84, 34.43, 47.11, 54.25, 171.20. ESIMS for $C_6H_{14}N_2O_2S$ [MH$^+$]: 178.25. Found: 178.2526.
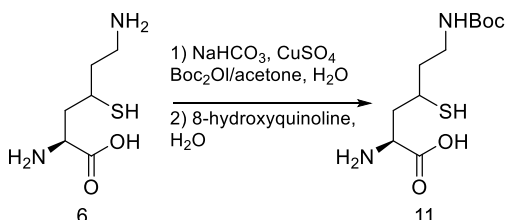
(2S)-2-amino-6-((*tert*-butoxycarbonyl)amino)-4-mercaptohexanoic acid (**11**): γ-mercaptolysine **6** (0.05 g, 0.28 mmol) was suspended in 0.22 mL of water. Sodium bicarbonate (0.09 g, 1.12 mmol) was added slowly. Then, CuSO$_4$ (0.02 g, 0.14 mmol) was added. Next, Boc anhydride (0.10 mL, 0.42 mmol) was added as a 1.51 M solution in acetone. The reaction was left stirring over 16 hours until a thick blue precipitate was formed. The precipitate was filtered under vacuum and rinsed with water. The powder was resuspended in 10 mL of water and 8-hydroxyquinoline (0.05g, 0.34 mmol) was added. The reaction mixture was left stirring until a green homogeneous suspension is detected. After vacuum filtration, the filtrate was

extracted with ethyl acetate (3 x 20 mL). The aqueous phase was distilled under vacuum and the white precipitate was dried at 40 °C, until the weight was stable. The carbamate was isolated (70%) and stored at 4 °C. ^{1}H NMR (D$_2$O, 400 MHz) δ 1.35 (s, 9H), 1.97, 2.09 (m, 4H), 2.89 (m, 1H), 3.30 (t, 6.39 Hz, 2H), 3.83 (s, 1H). ^{13}C NMR (D$_2$O, 100 MHz) δ 30.36, 31.84, 34.43, 47.11, 54.25, 64.59, 78.00, 156.07, 171.20. ESIMS for C$_{11}$H$_{22}$N$_2$O$_4$S [MH$^+$]: 278.37. Found: 280.3722.

Synthesis of δ-mercaptolysine

L-glutamic acid

Dimethyl (S)-2-tert-butoxycarbonylamino-pentanodioate: A solution of glutamic acid (2.627 g, 17.88 mmol) in 60 mL of methanol was cooled down to 0 °C. Then, trimethyl chlorosylane (10.00 mL, 78.68 mmol) was added slowly at this temperature. The reaction was then stirred over two days. Triethylamine (16.21 mL, 511.42 mmol) was added slowly to the reaction mixture, following by the addition of Boc anhydride (4.52 mL, 19.67 mmol) and left stirring over 16 hours. A second addition of Boc anhydride (2.26 mL, 0.01 mmol) was done and left stirring 16 more hours (*Rf: 0.6 Hex:EtOAc 80:20*). The methanol was removed under vacuum and then extracted with ethyl acetate (3 x 20 mL), dried over anhydrous sodium sulfate, the ethyl acetate was removed under vacuum and the product was isolated as an orange oil (97%). ^{1}H NMR (CDCl$_3$, 400 MHz) δ 1.39 (s, 9H), 1.91 (dt, 14.63, 8.02 Hz, 1H), 2.15 (m, 1H), 2.37 (m, 2H), 3.63 (s, 3H), 3.70 (s, 3H), 4.28 (dd, 12.96, 7.74 Hz, 1H) 5.12 (dd, 57.87, 7.56 Hz, 1H). ^{13}C NMR (CDCl$_3$, 100 MHz) δ 27.76, 29.54, 30.71, 51.27, 51.91, 52.36, 79.43, 154.85, 172.09, 172.64. ESIMS for C$_{12}$H$_{21}$NO$_6$ [MH+]: 276.29. Found: 276.2983.

Dimethyl (2S)-2-((tert-butoxy)-N-[(tert-butyl)oxy-carbonyl]carbonylamino)-pentanedioate **(C)**: A solution of dimethyl (S)-2-tert-butoxycarbonylamino-pentanodioate (4.43 g, 16.07 mmol) of was dissolved in acetonitrile (62.00 mL).

DMAP (0.39 g, 3.21 mmol) was then added to the reaction mixture, followed by Boc anhydride (4.10 mL, 17.68 mmol). After 2 hours of stirring, a second addition of Boc anhydride (2.05 mL, 8.84 mmol) was realized and then left stirring over 16 hours. The reaction mixture was extracted with ethyl acetate (3 x 30 mL) and washed with water (20 mL) and brine (20 mL); dried over anhydrous sodium sulfate and the ethyl acetate was removed under vacuum. The orange oil isolated was purified by column chromatography (*Rf: 0.75 Hex:EtOAc 80:20*) and a crystalline oil was isolated (90%). ^{1}H NMR (CDCl$_3$, 400 MHz) δ 1.47 (s, 18H), 2.17 (m, 1H), 2.39 (m, 2H), 2.46 (m, 1H), 3.65 (s, 3H), 3.69 (s, 3H), 4.93 (dd, 9.60, 4.83 Hz, 1H). ^{13}C NMR (CDCl$_3$, 100 MHz) δ 25.09, 27.85, 30.51, 51.56, 52.12, 57.24, 83.20, 151.81, 170.71, 172.99. ESIMS for C$_{17}$H$_{29}$NO$_8$ [MH$^+$]: 376.41. Found: 376.4143.
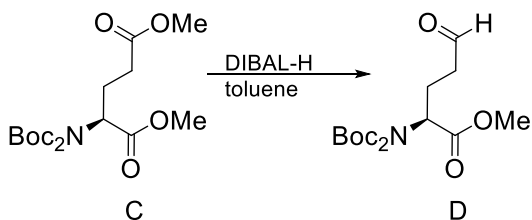
Methyl (2S)-2-((tert-butoxy)-N-[(*tert*-butoxyl)oxycarbonyl]-carbonylamino)-5-oxopentanoate (**D**): A solution of 5.083 g of **C** (13.53 mmol) in freshly distilled and dry toluene (135 mL), was flushed with argon for 30 minutes. Then, it was cooled down to -78 °C (dry ice - acetone bath). Then, DIBAL-H (28.50 mL, 1.2 M in toluene, 33.83 mmol) was added slowly and dropwise via syringe for over 15 minutes. The reaction was left stirring for 4 hours at -78 °C until no further formation of aldehyde was detected by TLC (*Rf: 0.45 Hex:EtOAc 75:25*). At -78 °C, the reaction was then quenched with 3 volumes of methanol, and left stirring until it reached -20 °C. The reaction mixture was stirred vigorously with 5 volumes of a saturated solution of Rochelle's salt for over an hour. Finally, the reaction mixture was extracted with ether (3 x 50 mL). The ethereal phase was washed with brine (1 x 30 mL), dried over anhydrous sodium sulfate and the ether was removed under vacuum. After column chromatography, a colorless oil was isolated (*Rf: 0.4; 80:20 Hexane/Ethyl Acetate,* 80%). ^{1}H NMR (CDCl$_3$, 400 MHz) δ 1.45 (s, 18H), 2.12 (m, 1H), 2.48 (m, 3H), 3.68 (s, 3H), 4.86 (dd, 12.96, 7.74 Hz, 1H) 9.72 (t, 57.87, 7.56 Hz,

1H). 13C NMR (CDCl3, 100 MHz) δ 22.44, 27.69, 40.71, 52.20, 57.22, 83.41, 151.88, 170.59, 200.67. ESIMS for $C_{16}H_{27}NO_7$ [MH$^+$]: 346.39. Found: 346.3882.

Methyl (5R,S)-hydroxy-6-nitro-(2S)-2-((*tert*-butoxy)-N-((*tert*-butoxyloxy)carbonyl)-carbonylamino)-hexanoate: Aldehyde **D** (3.37 g, 9.76 mmol) was dissolved in nitromethane (11.00 mL, 205.01 mmol). The reaction mixture was then cooled down to 0 °C and flushed with argon for 30 minutes. Then, TBAF (4.89 mL (1M in THF), 4.88 mmol) was added via syringe. The reaction was left stirring for 15 minutes. The reaction mixture was purified by column chromatography (*Rf: 0.3 Hex:EtOAc 70:30, 90%*) and isolated as a clear yellow oil. ^{1}H NMR (CDCl$_3$, 400 MHz) δ 1.47 (s, 18H), 1.57 (m, 2H), 2.00 (m, 1H), 2.31 (m, 1H), 2.94 (br d, 11.92 Hz, 1H), 3.70 (s, 3H), 4.38 (m, 3H), 4.86 (m, 1H). ^{13}C NMR (CDCl$_3$, 100 MHz) δ 26.11, 27.95, 30.41, 52.34, 57.38, 67.88, 80.44, 83.53, 152.09, 171.20. ESIMS for $C_{17}H_{30}N_2O_9$ [MH$^+$]: 407.19. Found: 407.1952.

Methyl 6-nitro-(5Z/E)-(2S)-2-((*tert*-butoxy)-N-[(*tert*-butyloxy)carbonyl]-carbonylamino)-hexa-5-enoate (**7**): The previous nitro alcohol (3.21 g, 7.89 mmol) was dissolved in 40 mL of dry diethyl ether. The solution was cooled down to 0 °C and flushed with argon for 10 minutes. At this temperature, acetic anhydride (0.89 mL, 9.74 mmol) was added via syringe followed by a solution of DMAP (0.39 g in 2 mL of ether, 3.16 mmol). The reaction was left stirring for a period of 2 hours (*Rf: 0.4, Hex:EtOAc 80:20*). The reaction mixture was quenched with 51 mL of saturated NH$_4$Cl and extracted with ether (3 x 40 mL). The ethereal phase was finally washed with brine (1 x 20mL) and dried over anhydrous sodium sulfate. The ether was removed under reduced pressure. The Michael acceptor (**7**) was purified by column chromatography, 90%. ^{1}H NMR (CDCl$_3$, 400 MHz) δ 1.47 (s, 18H), 2.06, 2.35 (m, 4H), 3.70 (s, 3H), 4.84 (q, 4.62 Hz, 1H), 5.44 (m, 1H, Z-isomer), 6.98 (m, 1H, E-isomer), 7.23 (m, 1H, E-isomer). ^{13}C NMR (CDCl$_3$, 100 MHz) δ 25.70, 27.94, 28.11,

52.38, 57.30, 83.67, 140.00 (E-isomer), 141.26 (Z-isomer), 152.09, 170.57. ESIMS for $C_{17}H_{28}N_2O_8$ [MH$^+$]: 389.18. Found: 389.1846.
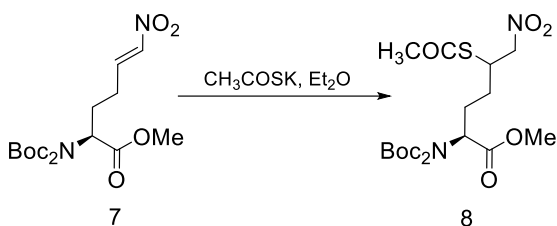
Methyl (5R,S)-(acetylthio)-2-((ditert-butoxycarbonyl)amino)-6-nitrohexanoate **(8)**: The nitroalkene (**7**) (2.39 g, 6.15 mmol) was dissolved in 31 mL of ether. Then, the system was cooled down to -15 °C and flushed with argon for 15 minutes. Potassium acetate (1.42 g, 12.3 mmol) was dissolved in 5 mL of ether and added dropwise via syringe to the nitroalkene solution. After three hours, the reaction was diluted with ether (20 mL). The reaction mixture was washed with water (2 x 20 mL) and the ether phase was dried over anhydrous sodium sulfate. The ether was removed under vacuum and the thioester was isolated as an orange oil after column chromatography (*Rf: 0.5 Hex:EtOAc 65:35,* 95%), and stored at 4 °C. ^{1}H NMR (CDCl$_3$, 400 MHz) δ 1.47 (s, 18H), 1.66 (m, 4H), 2.34 (s, 3H), 3.29 (m, 1H), 3.70 (s, 3H), 4.04 (m, 1H), 4.54 (m, 1H), 4.82 (ddd, 5.35, 9.21, 14.29 Hz, 1H). ^{13}C NMR (CDCl$_3$, 100 MHz) 26.88, 27.39, 27.95, 30.65, 40.63, 40.93, 52.31, 57.14, 65.85, 77.20, 83.52, 151.98, 170.93, 194.08. ESIMS for $C_{19}H_{32}N_2O_9S$ [MH$^+$]: 465.18. Found: 465.1829.
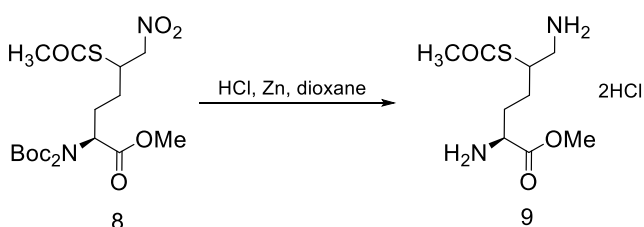
Methyl (2S)-5-(acetylthio)-2,6-diaminohexanoacetate (**9**): The nitrothioester (**8**) (0.79 g, 1.70 mmol) was dissolved in HCl (10 mL, 12 N in dioxane). Then, Zn powder (2.11 g, 32.31 mmol) was added slowly in small portions since H$_2$ was developed. The reaction is monitored by TLC until no starting material was observed and a very polar product was produced (*Rf: 0.15 Hex:EtOAc 65:35*). The reaction mixture was then filtered and rinsed with 10 mL of ethanol. Finally the ethanol was removed

under vacuum and the orange oil (70%) was stored at 4°C. ESIMS for $C_9H_{18}N_2O_3S\cdot 2HCl$ [MH$^+$]: 309.10. Found 309.1230.
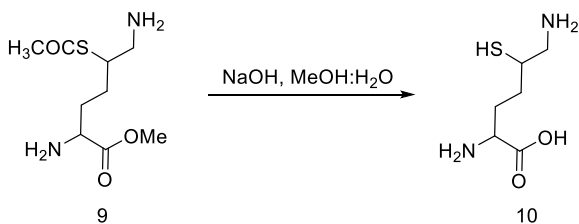
(2S)-(2,6)-diamino-5-mercaptohexanoic acid (**10**): Thioester (**9**) (0.72 g, 1.27 mmol) was dissolved in 15 mL of MeOH:H_2O (2:1). To the solution, 1N NaOH (5.73 mL, 5.73 mmol) was added via syringe over a period of ten minutes under argon atmosphere. The reaction mixture was left stirring six hours at room temperature. The solvent was removed under vacuum and the amino acid was triturated with ether (15 mL, 70%). ^{1}H NMR (DMSO-d_6, 400 MHz) δ 1.41 (m, 2H), 1.77 (d, 5.35 Hz, 2H), 2.89 (m, 1H), 3.21 (m, 2H), 3.83 (s, 1H), 8.41 (brs, 2H). ^{13}C NMR (DMSO-d_6, 100 MHz) δ 27.67, 30.77, 35.85, 47.99, 51.56, 170.98. ESIMS for $C_6H_{14}N_2O_2S$ [MH$^+$]: 178.25. Found: 180.3811.
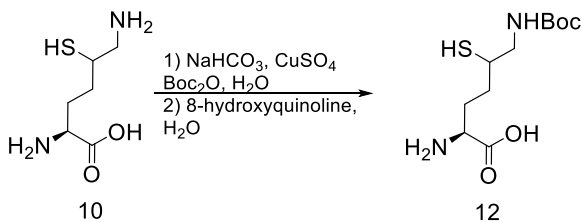
(2S)-2-amino-6-((*tert*-butoxycarbonyl)amino)-5-mercaptohexanoic acid (**12**): δ-mercaptolysine **10** (0.05 g, 0.28 mmol) was suspended in 0.22 mL of water. Sodium bicarbonate (0.09 g, 1.12 mmol) was added slowly. Then, CuSO$_4$ (0.02 g, 0.14 mmol) was added. Next, Boc anhydride (0.10 mL, 0.42 mmol) was added as a 1.51 M solution in acetone. The reaction was left stirring over 16 hours until a thick blue precipitate was formed. The precipitate was filtered under vacuum and rinsed with water. The powder was resuspended in 10 mL of water and 8-hydroxyquinoline (0.05g, 0.34 mmol) was added. The reaction mixture was left stirring until a green homogeneous suspension was detected. After vacuum filtration, the filtrate was extracted with ethyl acetate (3 x 20 mL). The aqueous phase was distilled under vacuum and the white precipitate was dried at 40 °C, until the weight was stable. The

carbamate was isolated (70%) and stored at 4 °C. ^{1}H NMR (DMSO-d$_6$, 400 MHz) δ 1.25 (m, 2H), 1.35 (s, 9H), 1.62 (m, 2H), 2.88 (dd, 5.93, 11.98 Hz, 2H), 3.16 (t, 5.71 Hz, 1H), 3.83 (s, 1H), 6.72 (brt, 5.57 Hz, 1H). ^{13}C NMR (DMSO-d$_6$, 100 MHz) δ 22.57, 29.48, 30.74, 47.99, 54.25, 78.00, 155.86, 170.70. ESIMS for C$_{11}$H$_{22}$N$_2$O$_4$S [MH$^+$]: 279.37. Found: 280.3713.

Protein expression tests

General procedure:

Transformation:

In a sterile environment 1.5 μL of plasmids in 15 μL of BL21DE3 E. coli cells were incubated 30 minutes at 0 °C. Then, it was incubated for 90 seconds at 42 °C and finally one more minute at 0 °C. 0.50 mL of LB media were added and left growing at 37 °C for one hour. Plated over LB agar with the corresponding antibiotic system (Table 9), and left growing at 37 °C overnight.

Induction:

A single colony was picked and grew in 10 mL of LB media with the corresponding antibiotic system for 6 hours at 37 °C. Then, 1 mL of culture was induced with IPTG/arabinose (final concentration 1 mM), and the amino acid (final concentration 2 mM) and left growing for 16 h at 30 °C.

SDS-PAGE:

The media was removed by centrifugation and the pellet was resuspended in 0.20 mL of sodium phosphate buffer (pH 6.8) twice, discarding the supernatant after centrifugation. Finally the pellet was resuspended in 0.10 mL of water.

An aliquot of 10 μL was mixed with 10 μL of 2x SDS loading buffer (containing b-mercaptoethanol) and 20 μL of water. The sample was heated up for 3 minutes at 90 °C. An aliquot of 10 μL was loaded on a 12% acrylamide gel with a standardized molecular weight marker for reference. The gel was run for a period of 40 minutes at 200 mV in a Tris pH 8.0/glycine buffer, stained with a methanolic Coomasie solution for one hour and destained with 10% acetic acid.

Plasmid systems	Antibiotic Resistance	Monosaccharide inducer
pPyls Y3884F/ sfGFP	Ampicilin/Chloramphenicol	IPTG
pBKev17/sfGFP	Kanamicin/Tetracyclin	Arabinose
pSupInterm+/pTrc HisA Y151 A sfGFP	Kanamicin/Chloramphenicol	Arabinose
pSupONB N436A/ pTrc HisA Y151 A sfGFP	Ampicilin/Chloramphenicol	IPTG
pSupONBLys/ pTrc HisA Y151 A sfGFP	Amplicilin/Chloramphenicol	IPTG
pSupInterm/pTrc HisA Y151 A sfGFP	Kanamicin/Chloramphenicol	Arabinose

Table 8. Protein expression test plasmid systems.

Native Chemical Ligation

Ub expression and purification:

Ub (Ub) gene inserted in the pET3a (ampicillin-resistant) vector was expressed in BL21(DE3)pJY2 *E. coli* strain cells (chloramphenicol-resistant) and grew in auto-inducing media at 37 °C in the presence of glucose and lactose. Per liter of growth, the lysis of cells was performed by sonication in the presence of 50 mM Tris (pH 7.6), 0.4mg/mL Lysozyme, 1 mM phenylmethylsulfonyl fluoride (PMSF), 50 μM tosyllysylchloromethyl ketone (TLCK), 2.5 μg soybean trypsin inhibitor (STI) and 2.5 μg leupeptin. After centrifugation (25000 rpm), 0.50 mL of 70% perchloric acid were added to the cold supernatant. The milky solution was then centrifuged at 25000 rpm and dialyzed against 50 mM ammonium acetate for 12 hours at 4 °C. Two buffer solutions were prepared for cation exchange chromatography (SPFF column, GE Life Sciences): Buffer A contained 50 mM ammonium acetate and buffer B contained 50 mM ammonium acetate and 1 M NaCl. A gradient from 0 to 40% B was used over 25 column volumes. Ub was eluted at 10-15% B. The corresponding fractions were collected, concentrated down, and buffer exchanged with 20 mM sodium phosphate (pH 6.8) and the protein was stored at -20 °C.

Ub thioesterification:

Ub (5 mg, 0.58 μmol) was buffer exchanged with 20 mM sodium phosphate buffer (pH 8.0) to a final volume of 0.50 mL. At 0 °C, the following solutions were added to the Ub solution: MesNa (100 μL, 1M, 100 μmol), ATP (100 μL, 0.1 M, 10 μmol), MgCl$_2$ (10 μL, 1M, 10 μmol). Finally, E1 enzyme was added to a final concentration of 1.2 μM, and the volume was adjusted to 1 mL of reaction. The system was incubated at 37 °C for six hours. Then, it was quenched with 20 μL of glacial acetic acid and buffer exchanged with MiliQ water grade, lyophilized and stored at -20 °C.

Native Chemical Ligation (NCL) with Ub

Ub-thioester [UbCOSR, (10 mg, 1.38 μmol)] was dissolved in 1.8 mL of sodium phosphate buffer (200 mM, pH 7.4), buffer exchanged and stored at 4 °C. In parallel, the mercaptolysine (4 mg, 18.69 μmol) was dissolved in 1.5 mL of NCL buffer (200 mM sodium phosphate pH 7.4, 100 mM mercaptophenylacetic acid, 60 mM TCEP). The solutions were combined and the pH was adjusted at 7.2 with HCl 5%. The ligation was left to proceed for 10 hours at 25 °C with gentle agitation. Ligation efficiency was 80%. The reaction was quenched by the addition of 1M TCEP dissolved in 4 M NaOH (10 μL). The reaction mixture was dialyzed against 200 mM sodium phosphate (pH 7.4) at 4 °C for at least 12 hours. A second dialysis was performed against 50 mM ammonium acetate (pH 4.5) supplemented with 5 mM dithiothreitol (DTT) at 4 °C for at least 12 hours. Finally, HP-cation exchange chromatography was used to separate Ub from Ub~(SH)K77 (HPSP column, GE Life Sciences) with the same gradient used for Ub purification. Ub~(SH)K77 fractions were pooled together and buffer exchanged three times with 20 mM sodium phosphate buffer (pH 6.8).

YUH1 hydrolysis

Ub~(SH)K77 (0.5 mg, 57.32 nmol) was buffer exchanged with 50 mM HEPES buffer (pH 8.0) to a final concentration of 29 mM. To the solution EDTA was added to a final concentration of 10 mM. Finally, YUH1 enzyme (5.0 μL, 230 mM) was added. The protein was incubated at 25 °C for 4 hours. The reaction was quenched with 10 μL of glacial acetic acid, and buffer exchanged with 20 mM sodium phosphate (pH 6.8)

C terminus labeling

Fluorescein-5-maleimide attachment

(δSH)K77Ub (2.0 mg, 229.28 nmol) was buffer exchanged with 50 mM HEPES buffer (pH 8.0). Then, a stoichiometric amount of fluorescein maleimide (229.28 nmol) was added. The reaction mixture was incubated at 25 °C with gentle stirring for a period of 10 hours. The protein was buffer exchanged twice with 20 mM sodium phosphate buffer (pH 6.8).

Spin labeling of δSHK77Ub

^{15}N- Ub(δSH)K77 (1.76 mg, 200 nmol) was buffer exchanged with 50 mM HEPES (pH 8.0) supplemented with 5 mM TCEP to a final concentration of 0.50 mM and incubated for four hours. Then, buffer exchanged twice with 50 mM HEPES (pH 8.0), to a final concentration of 0.20 mM. Next, MTSL [(1-oxyl-2,2,5,5,-tetramethylpyrroline-3-methyl) methanethiosulfonate] (15 μL, 40 mM, 600 nmol) and incubated for at least 16 hours in the dark. Finally buffer exchanged with 20 mM sodium phosphate (pH 6.8) three times. After NMR-HSQC measurements the MTSL was reduced by the addition of ascorbate (30 μL, 100 mM, 1.20 mmol). Finally, the MTSL could be removed by incubation with TCEP, and buffer exchange with 20 mM sodium phosphate (pH 6.8).

Section 2. Silver Assisted Ligation Reaction

Synthesis of N-ε-AllocLysine

Lysine (0.092 mg, 0.51 mmol) was suspended in 0.40 mL of water. Sodium bicarbonate (0.172 g, 2.05 mmol) was added slowly. Then, $CuSO_4$ (0.04 g, 0.26 mmol) was added. Next, Alloc chloride (0.08 mL, 0.0.77 mmol) was added as a 1.51 M solution in acetone. The reaction was left stirring over 16 hours until a thick blue precipitate was formed. The precipitate was filtered under vacuum and rinsed with

water. The powder was resuspended in 5 mL of water and 8-hydroxyquinoline (0.09 g, 0.61 mmol) was added. The reaction mixture was left stirring until a green homogeneous suspension was detected. After vacuum filtration, the filtrate was extracted with ethyl acetate (3 x 20 mL). The aqueous phase was distilled under vacuum and the white precipitate was dried at 40 °C, until the weight was stable. The carbamate was isolated (85%) and stored at 4 °C. ^{1}H NMR (CD$_3$OD, 400 MHz) δ 1.37 (m, 2H), 1.58, 1.72 (m, 4H), 2.96, 3.12 (m, 3H), 4.46 (s, 2H), 5.25, 5.30 (dq, 1.68, 17.37Hz, 2H), 5.92 (ddd, 5.34, 10.57, 22.34 Hz, 1H), 7.22 (t, 5.13 Hz, 1H), 7.6 (brs, 2H). ^{13}C NMR (CD$_3$OD, 100 MHz) δ 22.63, 29.70, 31.11, 40.54, 55.15, 65.29, 116.44, 133.66, 157.92, 173.40. ESIMS for C$_{10}$H$_{18}$N$_2$O$_4$ [MH$^+$]: 230.37. Found: 230.3823.

K48Alloc Ub

2.5 μL of plasmids pTXB1K48TAGUb and pSup-PylT-PylS Y384F were transformed into 50 μL BL21DE3 *E. coli* cells were incubated 30 minutes at 0 °C. Then, it was incubated for 90 seconds at 42 °C and finally one more minute at 0 °C. 0.50 mL of LB media were added and left growing at 37 °C for one hour. Plated over LB agar (ampicillin/chloramphenicol) and left growing at 37 °C overnight. A single colony was picked and grew in 20 mL of LB (ampicillin/chloramphenicol) for 6 hours at 37 °C. Then, 1 L of magic media (ampicillin/chloramphenicol) was inoculated in the presence of glucose, lactose, and N-ε-alloclysine (5 mM final concentration) and left growing for 16 h at 30 °C. Per liter of growth, the lysis of cells was performed by sonication in the presence of 20 mM HEPES (pH 8.0), 1 mM phenylmethylsulfonyl fluoride (PMSF), 0.25 μg soybean trypsin inhibitor (STI) and 0.25 μg leupeptin and 0.1% Triton X-100. After centrifugation (25000 rpm), the protein was purified by intein-mediated cleavage of a chitin binding domain (CBD) with 50 mM DTT, and by SEC chromatography. Finally, it was buffer exchanged in 20 mM sodium phosphate (pH 6.8), 10 mg of K48AllocUb were isolated and lyophilized for further reactions.

Ub thioesterification

Ub (5 mg, 0.58 µmol) was buffer exchanged with 20 mM sodium phosphate buffer (pH 8.0) to a final volume of 0.50 mL. At 0 °C, the following solutions were added to the Ub solution: MesNa (100 µL, 1M, 100 µmol), ATP (100 µL, 0.1 M, 10 µmol), $MgCl_2$ (10 µL, 1M, 10 µmol). Finally, E1 enzyme was added to a final concentration of 1.2 µM, and the volume was adjusted to 1 mL of reaction. The system was incubated at 37 °C for six hours. Then the reaction was quenched with 20 µL of glacial acetic acid and buffer exchanged with MiliQ water grade, lyophilized and stored at -20 °C.

K48BocUb

2.5 µL of plasmids pTXB1K48TAGUb and pSup-PylT-PylS Y384F were transformed into 50 µL BL21DE3 *E. coli* cells were incubated 30 minutes at 0 °C. Then, it was incubated for 90 seconds at 42 °C and finally one more minute at 0 °C. 0.50 mL of LB media were added and left growing at 37 °C for one hour. Plated over LB agar (ampicillin/chloramphenicol) and left growing at 37 °C overnight. A single colony was picked and grew in 20 mL of LB (ampicillin/chloramphenicol) for 6 hours at 37 °C. Then, 1 L of magic media (ampicillin/chloramphenicol) was inoculated in the presence of glucose, lactose, and N-ε-Boc lysine (5 mM final concentration) and left growing for 16 h at 30 °C. Per liter of growth, the lysis of cells was performed by sonication in the presence of 20 mM HEPES (pH 8.0), 1 mM phenylmethylsulfonyl fluoride (PMSF), 0.25 µg soybean trypsin inhibitor (STI) and 0.25 µg leupeptin and 0.1% Triton X-100. After centrifugation (25000 rpm), the protein was purified by intein-mediated cleavage of a chitin binding domain (CBD) with 50 mM DTT, and by SEC chromatography. Finally, it was buffer exchanged in 20 mM sodium phosphate (pH 6.8), 25 mg of K48BocUb were isolated and lyophilized for further reactions.

Orthogonal carbamate protection

a) (K48Alloc Ub)Boc$_{8-10}$.

The lyophilized protein was dissolved in DMSO to a final concentration of 2.40 mM. To the solution, 160 equivalents of DIEA were added. Following, 30 equivalents of SuOH (10mg/mL, in DMSO) and 30 equivalents of Boc$_2$O (neat) were added. The reaction was left in the rocket stirrer 6 hours at room temperature. Finally, the reaction mixture is ether precipitated as follows: The DMSO phase was separated in 100µL aliquots and to each one, 500 µL of cold ether were added. Then, vortexed vigorously for a minute and spun down in cold for 10 minutes at 12K rpm. This precipitation step was repeated two more times. A final ether precipitation is done with the extra addition of 100 µL of ACN. The protein falls as a whitish precipitate which was left air drying. Then, it was characterized by MS (ESI$^+$). The product is consisted to the incorporation of 8 to 10tert-butoxycarbonyl groups, [(K48Alloc)Boc$_{8-10}$]Ub.

b) (UbCOSR)Boc$_{8-10}$

The lyophilized UbCOSR was dissolved in DMSO to a final concentration of 2.40 mM. To the solution, 160 equivalents of DIEA and subsequently 30 equivalents of SuOH (10mg/mL) and Boc$_2$O (neat) were added. The reaction was left in the rocket stirrer 6 hours at room temperature. Finally the reaction mixture is ether precipitated as previously mentioned. Then it was left air drying and characterized by MS (ESI+). The product is consisted with the incorporation of 8 to 10 tert-butoxycarbonyl groups, [UbCOSR]Boc$_{8-10}$

c) (K48BocUb)Alloc$_8$

The lyophilized protein was dissolved in DMSO to a final concentration of 2.40 mM. To the solution, 160 equivalents of DIEA were added. Following, 30 equivalents of Alloc-SuOH (10mg/mL, in DMSO) were added. The reaction was left in the rocket stirrer 6 hours at room temperature. Finally, the reaction mixture is ether precipitated as follows: The DMSO phase was separated in 100µL aliquots and to each one, 500 µL of cold ether were added. Then, vortexed vigorously for a minute and spun down

in cold for 10 minutes at 12K rpm. This precipitation step was repeated two more times. A final ether precipitation is done with the extra addition of 100 µL of ACN. The protein falls as a whitish precipitate which was left air drying. Then, it was characterized by MS (ESI$^+$). The product is consisted to the incorporation of 8 allyloxycarbonyl groups, [(K48Boc)Alloc$_8$]Ub.

d) (UbCOSR)Alloc$_{8-10}$

The lyophilized UbCOSR was dissolved in DMSO to a final concentration of 2.40 mM. To the solution, 160 equivalents of DIEA and subsequently 30 equivalents of Alloc-SuOH (10mg/mL) were added. The reaction was left in the rocket stirrer 6 hours at room temperature. Finally the reaction mixture is ether precipitated as previously mentioned. Then it was left air drying and characterized by MS (ESI+). The product is consisted with the incorporation of 8 to 10 allyloxycarbonyl groups, [UbCOSR]Alloc$_{8-10}$

e) (K48Ub)Boc$_{8-10}$

(K48AllocUb)Boc$_8$ was solubilized in DMSO as the 60 % of the solvent. Then, water was added to complete the 100% towards a total concentration of the protein equal to 1.23 mM. To this solution, it was added 0.53 equivalents of RuCp* catalyst (10mg/mL in DMSO). Then, 5.33 equivalents of thiophenol are added to the reaction mixture which is exothermic and continue pipetting is needed to avoid the presence of a black precipitate. Finally, the reaction mixture is incubated at 50 °C for 4 hours and then at room temperature overnight. Ether precipitation follows after the incubation as previously mentioned. The product is characterized by MS (ESI+) and it shows the total cleaveage of the allyloxycarbonyl group as a (K48Ub)Boc$_{8-10}$

f) (K48Ub)Alloc$_{8-10}$

(K48Alloc)Boc$_8$Ub was solubilized in DMSO as the 60 % of the solvent. Then, 60% TFA was added to complete the 100% towards a total concentration of the protein equal to 1.23 mM. The reaction mixture is incubated at 25 °C for 4 hours and then at room temperature overnight. Ether precipitation follows after the incubation as previously mentioned. The product is characterized by MS (ESI+) and it shows the total cleaveage of the tert-butoxycarbonyl group as a (K48Ub)Alloc$_{8-10}$

Silver assisted ligation reaction

Both monomers were dissolved in DMSO to a final concentration of 2.00 mM. Then, 163 equivalents of DIEA were added. Following, the addition of 30 equivalents of SuOH (10mg/mL) was made. Then, 3 equivalents of $AgNO_3$ (10 mg/mL) were added and finally the (UbCOSR)Boc$_8$ or (UbCOSR)Alloc$_8$ was added to the reaction mixture. Then, left stirring at room temperature and covered from light for 48 hours. After this period of time, the dimer was precipitated as previously described.

Orthogonal cleavage reaction

Once the protein dimer was precipitated, the pellet was redissolved with a 60% TFA solution if the Boc group was going to be deprotected and left stirring at 4 C for 4 hours. In the case of the Alloc group, the protein dimer was redissolved in DMSO as the 60 % of the solvent. Then, water was added to complete the 100% toward a total concentration of the protein equal to 1.23 mM. To this solution, it was added 0.53 equivalents of RuCp* catalyst (10mg/mL in DMSO). Then, 5.33 equivalents of thiophenol are added to the reaction mixture which is exothermic and continue pipetting is needed to avoid the presence of a black precipitate. Finally, the reaction mixture is incubated at 50 °C for 4 hours and then at room temperature overnight. Ether precipitation follows after the incubation as previously mentioned. Once again, it was precipitated with ether thrice and finally renatured by dialysis and finally purified by cation exchange chromatography.

Section 3. K48Ub$_2$ Dimers

Natural enzymatic dimer

Proximal domain:

D77Ub mutant inserted in the pET3a (ampicillin-resistant) vector was expressed in BL21(DE3)pJY2 *E. coli* strain cells (chloramphenicol-resistant) and grew in auto-inducing media at 30 °C in the presence of glucose and lactose. Per liter of growth, the lysis of cells was performed by sonication in the presence of 50 mM Tris (pH

7.6), 0.4mg/mL Lysozyme, 1 mM phenylmethylsulfonyl fluoride (PMSF), 50 µM tosyllysylchloromethyl ketone (TLCK), 2.5 µg soybean trypsin inhibitor (STI) and 2.5 µg leupeptin. After centrifugation (25000 rpm), 0.50 mL of 70% perchloric acid were added to the cold supernatant. The milky solution was then centrifuged at 25000 rpm and dialyzed against 50 mM ammonium acetate for 12 hours at 4 °C. Two buffer solutions were prepared for cation exchange chromatography (SPFF column, GE Life Sciences): Buffer A contained 50 mM ammonium acetate and buffer B contained 50 mM ammonium acetate and 1 M NaCl. A gradient from 0 to 40% B was used over 25 column volumes. Ub was eluted at 10-15% B. The corresponding fractions were collected, concentrated down, and buffer exchanged with 20 mM sodium phosphate (pH 6.8), 130 mM NaCl and purified by SEC chromatography. Finally, D77Ub was buffer exchanged with 20 mM sodium phosphate (pH 6.8) and stored at -20 °C.

Distal domain

2.5 µL of plasmids pTXB1K48TAGUb and pSup-PylT-PylS Y384F were transformed into 50 µL BL21DE3 *E. coli* cells were incubated 30 minutes at 0 °C. Then, it was incubated for 90 seconds at 42 °C and finally one more minute at 0 °C. 0.50 mL of LB media were added and left growing at 37 °C for one hour. Plated over LB agar (ampicillin/chloramphenicol) and left growing at 37 °C overnight. A single colony was picked and grew in 20 mL of LB (ampicillin/chloramphenicol) for 6 hours at 37 °C. Then, 1 L of magic media (ampicillin/chloramphenicol) was inoculated in the presence of glucose, lactose, and N-ε-Boc lysine (5 mM final concentration) and left growing for 16 h at 30 °C. Per liter of growth, the lysis of cells was performed by sonication in the presence of 20 mM HEPES (pH 8.0), 1 mM phenylmethylsulfonyl fluoride (PMSF), 0.25 µg soybean trypsin inhibitor (STI) and 0.25 µg leupeptin and 0.1% Triton X-100. After centrifugation (25000 rpm), the protein was purified by intein-mediated cleavage of a chitin binding domain (CBD) with 50 mM DTT, and by SEC chromatography. Finally, it was buffer exchanged in 20 mM sodium phosphate (pH 6.8), 25 mg of K48AllocUb were isolated and lyophilized for further reactions.

Ligation reaction

In a 2 mL of reaction, 15 mg of monomer in 50 mM Tris buffer (pH 8.0) were added 400 μL of PBDM8, 40 μL of 0.1M ATP, 60 μL of 0.1 M TCEP, and 5 μL of E2-25K (250 mM). After mixed through E1 was added (final concentration 100 nM), and the reaction was left incubated at 37 C for 8 hours. The reaction was then quenched by the addition of 50 μL of glacial acetic acid, and purified by high-performance cation exchange chromatography with the same gradient used for Ub separation. The dimer eluted at 20-30% B. The dimer was then buffer exchanged to 50mM HEPES (pH 8.0) to hydrolyze with the YUH1 enzyme.

YUH1 hydrolysis

$K48Ub_2$ in HEPES buffer (pH 8.0) to a final concentration of 29 mM was added EDTA to a final concentration of 10 mM. Finally, YUH1 enzyme (5.0 μL, 230 mM) was added. The protein was incubated at 25 °C for 4 hours. The reaction was quenched with 10 μL of glacial acetic acid, and buffer exchanged with 2% TFA for Boc deprotection.

TFA cleavage

$K48Ub_2$ in 2% TFA was left at 4 °C for 6 hours. Then it was buffer exchanged to 50mM ammonium acetate and purified by high-performance cation exchange chromatography with the same gradient used for Ub, the dimer eluted at 20-30% B. Finally the dimer is buffer exchanged with 20 mM sodium phosphate (pH 6.8) and stored at -20 °C.

Thiol-ene click chemistry

Proximal domain

K48C Ub mutant inserted in the pET3a (ampicillin-resistant) vector was expressed in BL21(DE3)pJY2 *E. coli* strain cells (chloramphenicol-resistant) and grew in auto-inducing media at 30 °C in the presence of glucose and lactose. Per liter of growth,

the lysis of cells was performed by sonication in the presence of 50 mM Tris (pH 7.6), 0.4mg/mL Lysozyme, 1 mM phenylmethylsulfonyl fluoride (PMSF), 50 μM tosyllysylchloromethyl ketone (TLCK), 2.5 μg soybean trypsin inhibitor (STI) and 2.5 μg leupeptin. After centrifugation (25000 rpm), the supernatant was heated up to 65 °C for 15 minutes and then cooled down at 0 °C for 10 minutes. After centrifugation (25000 rpm), the supernatant was dialyzed against 50 mM ammonium acetate for 12 hours at 4 °C. A final centrifugation round (25000 rpm), the protein was purified by cation exchange chromatography (SPFF column, GE Life Sciences): Buffer A contained 50 mM ammonium acetate and buffer B contained 50 mM ammonium acetate and 1 M NaCl. A gradient from 0 to 40% B was used over 25 column volumes. K48C Ub was eluted at 10-15% B. The corresponding fractions were collected, concentrated down, and buffer exchanged with 20 mM sodium phosphate (pH 6.8), 130 mM NaCl and purified by SEC chromatography. Finally, K48C Ub was buffer exchanged with 20 mM sodium phosphate (pH 6.8) and stored at -20 °C.

Distal domain

D77Ub

D77Ub mutant inserted in the pET3a (ampicillin-resistant) vector was expressed in BL21(DE3)pJY2 *E. coli* strain cells (chloramphenicol-resistant) and grew in auto-inducing media at 30 °C in the presence of glucose and lactose. Per liter of growth, the lysis of cells was performed by sonication in the presence of 50 mM Tris (pH 7.6), 0.4mg/mL Lysozyme, 1 mM phenylmethylsulfonyl fluoride (PMSF), 50 μM tosyllysylchloromethyl ketone (TLCK), 2.5 μg soybean trypsin inhibitor (STI) and 2.5 μg leupeptin. After centrifugation (25000 rpm), 0.50 mL of 70% perchloric acid were added to the cold supernatant. The milky solution was then centrifuged at 25000 rpm and dialyzed against 50 mM ammonium acetate for 12 hours at 4 °C. Two buffer solutions were prepared for cation exchange chromatography (SPFF column, GE Life Sciences): Buffer A contained 50 mM ammonium acetate and buffer B contained 50 mM ammonium acetate and 1 M NaCl. A gradient from 0 to 40% B was used over 25 column volumes. Ub was eluted at 10-15% B. The corresponding fractions were

collected, concentrated down, and buffer exchanged with 20 mM sodium phosphate (pH 6.8), 130 mM NaCl and purified by SEC chromatography. Finally, D77Ub was buffer exchanged with 20 mM sodium phosphate (pH 6.8) and stored at -20 °C.

UbAA

D77Ub was buffer exchanged into 50 mM HEPES (pH 8.0) to a final concentration of 570 mM. A standard 5 mg reaction of 1 mL was set up with a master mix of 70% water, 30% DMSO, 72 μL 500 mM HEPES (pH 8.0), 7.2 μL of 100 mM EDTA (pH 8.0) and 22 μL of allylamine. Then 5 uL (230 mM) of YUH1 enzyme were added and the reaction was left incubated for 2 h at room temperature. Then 10 uL of hydrochloric acid were added and buffer exchanged to 50 mM ammonium acetate. The protein was purified by high-performance cation chromatography using the same gradient to purify Ub. UbAA was isolated in 50% yield and buffer exchanged to 50 mM NaOAc pH 5.1 for free radical reaction.

The Ub recovered from the reaction was buffer exchanged into 20 mM sodium phosphate (pH 8.0) for thioesterification as previously described. The thioester was then left reacting with 30 equivalents of allylamine for 48 hours. Finally UbAA was purified by high-performance cation chromatography as previously described and buffer exchanged in NaOAc (pH 5.1) for free radical reaction.

Ligation reaction

To a standard reaction of 1mL final volume, 1 mg of monomers and V-50 (final concentration of 5 mM) were mixed and incubated at 37 °C for a period of 10 hours. Then, the reaction mixture was buffer exchanged into 50 mM of ammonium acetate and purified by high-performance cation chromatography. The dimer eluted at ~25-30% B. Finally the dimer was buffer exchanged into 20 mM sodium phosphate (pH 6.8) and stored at -20 °C.

Intermolecular disulfide exchange

Proximal domain

K48C Ub mutant inserted in the pET3a (ampicillin-resistant) vector was expressed in BL21(DE3)pJY2 *E. coli* strain cells (chloramphenicol-resistant) and grew in auto-inducing media at 30 °C in the presence of glucose and lactose. Per liter of growth, the lysis of cells was performed by sonication in the presence of 50 mM Tris (pH 7.6), 0.4mg/mL Lysozyme, 1 mM phenylmethylsulfonyl fluoride (PMSF), 50 μM tosyllysylchloromethyl ketone (TLCK), 2.5 μg soybean trypsin inhibitor (STI) and 2.5 μg leupeptin. After centrifugation (25000 rpm), the supernatant was heated up to 65 °C for 15 minutes and then cooled down at 0 °C for 10 minutes. After centrifugation (25000 rpm), the supernatant was dialyzed against 50 mM ammonium acetate for 12 hours at 4 °C. A final centrifugation round (25000 rpm), the protein was purified by cation exchange chromatography (SPFF column, GE Life Sciences): Buffer A contained 50 mM ammonium acetate and buffer B contained 50 mM ammonium acetate and 1 M NaCl. A gradient from 0 to 40% B was used over 25 column volumes. K48C Ub was eluted at 10-15% B. The corresponding fractions were collected, concentrated down, and buffer exchanged with 20 mM sodium phosphate (pH 6.8), 130 mM NaCl and purified by SEC chromatography. Finally, K48C Ub was buffer exchanged with 20 mM sodium phosphate (pH 6.8) and stored at -20 °C.

Distal domain

2.5 μL of pTXB1Ub plasmid was transformed into 50 μL BL21DE3 *E. coli* cells were incubated 30 minutes at 0 °C. Then, it was incubated for 90 seconds at 42 °C and finally one more minute at 0 °C. 0.50 mL of LB media were added and left growing at 37 °C for one hour. Plated over LB agar (ampicillin/chloramphenicol) and left growing at 37 °C overnight. A single colony was picked and grew in 20 mL of LB (ampicillin/chloramphenicol) for 6 hours at 37 °C. Then, 1 L of magic media (ampicillin/chloramphenicol) was inoculated in the presence of glucose and lactose, and left growing for 16 h at 30 °C. Per liter of growth, the lysis of cells was

performed by sonication in the presence of 20 mM HEPES (pH 8.0), 1 mM phenylmethylsulfonyl fluoride (PMSF), 0.25 µg soybean trypsin inhibitor (STI) and 0.25 µg leupeptin and 0.1% Triton X-100. After centrifugation (25000 rpm), the protein was purified by intein-mediated cleavage of a chitin binding domain (CBD) with 100 mM cysteamine (pH 8.0) and by SEC chromatography. Finally, it was buffer exchanged in 50 mM HEPES (pH 8.0).

Ligation reaction:

The distal domain was incubated with 5 mM diamide for a period of approximately 16 hours to create a linear disulfide dimer. The reaction was considered complete when no monomer was detected by SDS-PAGE. A 1.5 mL of 5 mg standard reaction was set up by addition of the proximal domain and diamide to a final concentration of 10 mM. The reaction was left incubating at room temperature for 24 hours. The reaction was quenched with 15 µL of acetic acid. The dimer was buffer exchanged with 50 mM ammonium acetate for high-performance cation exchange chromatography using the same gradient to purify Ub. The dimer was eluted ~25-30% of B. Finally the dimer was buffer exchanged into 20 mM sodium phosphate (pH 6.8) and stored at -20 °C.

Thiolysine enzymatic dimer

Proximal domain

K48CD77Ub

K48CD77Ub mutant inserted in the pET3a (ampicillin-resistant) vector was expressed in BL21(DE3)pJY2 *E. coli* strain cells (chloramphenicol-resistant) and grew in auto-inducing media at 30 °C in the presence of glucose and lactose. Per liter of growth, the lysis of cells was performed by sonication in the presence of 50 mM Tris (pH 7.6), 0.4mg/mL Lysozyme, 1 mM phenylmethylsulfonyl fluoride (PMSF), 50 µM tosyllysylchloromethyl ketone (TLCK), 2.5 µg soybean trypsin inhibitor (STI) and 2.5 µg leupeptin. After centrifugation (25000 rpm), the supernatant was heated up to 65 °C for 15 minutes and then cooled down at 0 °C for 10 minutes. After centrifugation (25000 rpm), the supernatant was dialyzed against 50 mM ammonium acetate, 5 mM TCEP for 12 hours at 4 °C. A final centrifugation round (25000 rpm), the protein was purified by cation exchange chromatography (SPFF column, GE Life Sciences): Buffer A contained 50 mM ammonium acetate, 5 mM TCEP and buffer B contained 50 mM ammonium acetate, 5 Mm TCEP and 1 M NaCl. A gradient from 0 to 40% B was used over 25 column volumes. K48CD77Ub was eluted at 10-15% B. The corresponding fractions were collected, concentrated down, and buffer exchanged with 20 mM sodium phosphate (pH 6.8), 5 Mm TCEP, and 130 mM NaCl and purified by SEC chromatography. Finally, K48CD77Ub was buffer exchanged with 50 mM Tris buffer (pH 8.0).

tK48D77Ub

A standard 1.5 mL reaction of 5 mg protein was incubated with 2-bromoethylamine hydrobromide (10 mM in Tris buffer pH 8.0 and 5 mM TCEP) at 42 °C for 24 hours. The reaction was quenched with 20 mL of acetic acid and dialyzed against 50 mM ammonium acetate, 5 mM TCEP. The tK48CD77Ub was purified by high-performance cation chromatography using the same gradient for Ub. The protein eluted at 13-15% of B. The protein was then buffer exchanged to 20 mM sodium phosphate buffer (pH 8.0).

Distal domain

2.5 µL of plasmids pTXB1K48TAGUb and pSup-PylT-PylS Y384F were transformed into 50 µL BL21DE3 *E. coli* cells were incubated 30 minutes at 0 °C. Then, it was incubated for 90 seconds at 42 °C and finally one more minute at 0 °C. 0.50 mL of LB media were added and left growing at 37 °C for one hour. Plated over LB agar (ampicillin/chloramphenicol) and left growing at 37 °C overnight. A single colony was picked and grew in 20 mL of LB (ampicillin/chloramphenicol) for 6 hours at 37 °C. Then, 1 L of magic media (ampicillin/chloramphenicol) was inoculated in the presence of glucose, lactose, and N-ε-Boc lysine (5 mM final concentration) and left growing for 16 h at 30 °C. Per liter of growth, the lysis of cells was performed by sonication in the presence of 20 mM HEPES (pH 8.0), 1 mM phenylmethylsulfonyl fluoride (PMSF), 0.25 µg soybean trypsin inhibitor (STI) and 0.25 µg leupeptin and 0.1% Triton X-100. After centrifugation (25000 rpm), the protein was purified by intein-mediated cleavage of a chitin binding domain (CBD) with 50 mM DTT, and by SEC chromatography. Finally, it was buffer exchanged in 20 mM sodium phosphate (pH 6.8), 25 mg of K48AllocUb were isolated and lyophilized for further reactions.

Ligation reaction

In a 2 mL of reaction, 15 mg of monomer in 50 mM Tris buffer (pH 8.0) were added 400 µL of PBDM8, 40 µL of 0.1M ATP, 60 µL of 0.1 M TCEP, and 5 µL of E2-25K (250 mM). After mixed through E1 was added (final concentration 100 nM), and the reaction was left incubated at 37 C for 8 hours. The reaction was then quenched by the addition of 50 µL of glacial acetic acid, and purified by high-performance cation exchange chromatography with the same gradient used for Ub separation. The dimer eluted at 20-30% B. The dimer was then buffer exchanged to 50mM HEPES (pH 8.0) to hydrolyze with the YUH1 enzyme.

YUH1 hydrolysis

K48Ub$_2$ in HEPES buffer (pH 8.0) to a final concentration of 29 mM was added EDTA to a final concentration of 10 mM. Finally, YUH1 enzyme (5.0 μL, 230 mM) was added. The protein was incubated at 25 °C for 4 hours. The reaction was quenched with 10 μL of glacial acetic acid, and buffer exchanged with 2% TFA for Boc deprotection.

TFA cleavage

K48Ub$_2$ in 2% TFA was left at 4 °C for 6 hours. Then it was buffer exchanged to 50mM ammonium acetate and purified by high-performance cation exchange chromatography with the same gradient used for Ub, the dimer eluted at 20-30% B. Finally the dimer is buffer exchanged with 20 mM sodium phosphate (pH 6.8) and stored at -20 °C.

Chapter 6: Conclusions

The post-translational modification of proteins is a system cells have to achieve and address signals in almost every pathway. The regulated protein degradation is essential for cell survival because it is necessary to respond immediately to environmental stimuli. Of these post-translational modifications ubiquitination is catalogued as reversible and it triggers different signaling when Ub is attached to a target protein. Therefore, understanding ubiquitination is a growing field and nowadays the application of chemical reactions allows the isolation of quantitative materials for structural studies. In this dissertation it was explored first the outcome of attaching a mercaptolysine residue by native chemical ligation at the C terminus of Ub keeping the nucleophilic sulfhydryl group in order to map the surface of Ub when a paramagnetic label is attached. To address this first aim the synthesis of the mercaptolysines was achieved by a more efficient route where the conditions of reaction with thioacetate as a nucleophile to a Michael addition system allowed a global yield improvement. In particular, the thioacetate reacted as a Michael donor in the conjugate addition either by the classic addition elimination reaction when the Michael acceptor had an electrowithdrawing group or by free radical reaction when a thermal azo-water initiator such as V-50 was used; especially since the initiator has been primarily employed in polymerization reactions. Following the synthesis of the mercaptolysines the most important result was the outcome from the native chemical ligation procedure. Besides the successful attachment of the mercaptolysine residue under non-denaturing conditions for Ub as

well as with the mercaptolysine being in solution without additional protection to drive the reaction regioselectively, it was possible to characterize the S→N shift preference when two primary amines are present by NMR. An isopeptide bond was preferred with respect a peptide bond; in addition, the attachment of the mercaptolysines to Ub was reversible after incubation with a Ub carboxyl-terminal hydrolase such as YUH1, and no alteration of the Ub structure was detected. Finally, it was also studied the effect of UBDs when an extra charge is present in the Ub structure showing that the UBDs still interact with the hydrophobic patch of Ub and the affinities remained in the same order of magnitude sufficing a 1:1 model.

By the silver assisted ligation it was possible to conclude that the orthogonal protection and deprotection of each monomer have an impact on the reaction yield. The non-enzymatic ligation strategy was useful to suggest the mechanism of reaction when silver was used as a catalyst, and also how the ruthenium catalyst could coordinate with the allylic system and the amine in a regioselectively manner without the interference of other aromatic side chains present in the protein.

Third, the formation of ubiquitin dimers was approached by building an inter-ubiquitin linkage mimicking the isopeptide bond with two approaches, the classic disulfide exchange as well as the thiol-ene click reaction by thermal initiation in aqueous conditions. After assembling the dimeric units, they were studied by Nuclear Magnetic Resonance, in order to establish a conformational state profile which depends on the pH conditions and it was also noticed their 3D structures of each Ub units were intact. On one hand, the thiol-ene chemistry dimers were synthesized using a water-soluble azo-compound V-50 by thermal free radical chemistry. By NMR it

was detected their hydrophobic patches are in a relatively closed conformation however it is possible to suggest the presence of two conformational states. In addition, the disulfide dimers were achieved by redox chemistry with diamide when the reactive species were at pH 8.0 in HEPES buffer. The dimer interface have a similarity with the natural dimers by showing perturbations around the hydrophobic patch and around the isopeptidic bond mimic when they are at physiological pH (7.4). Finally, both dimers exhibited different conformational properties, especially the ability to form a closed state due to the pH-dependence of the Ub_2 conformation reinforcing the concept of a switch and open the question to test their binding properties towards UBDs.

Altogether, this dissertation is a compilation of chemical-biological strategies to isolate proteins *in vitro* homogeneously and in higher quantities and from the vast organic chemistry toolbox of reactions, the ones selected were efficient, and it was a starting point toward the applicability of building scaffolds for drug discovery screening when Ub can be used as a starting point to identify binding partners.

Bibliography

1.   Nalepa, G.; Rolfe, M.; Harper, J. W., *Nat. Rev. Drug Discov.,* **2006,** *5*, 596-613.
2.   Ciechanover, A.; Kwon, Y. T., *Exp. Mol. Med.,* **2015,** *47*, e147.
3.   Wang, X.; Terpstra, E. J., *J. Mol. Cell Cardiol.,* **2013,** *55*, 73-84.
4.   Ciechanover, A.; Stanhill, A., *Biochim. Biophys. Acta,* **2014,** *1843*, 86-96.
5.   Williams, A.; Sarkar, S.; Cuddon, P.; Ttofi, E. K.; Saiki, S.; Siddiqi, F. H.; Jahreiss, L.; Fleming, A.; Pask, D.; Goldsmith, P.; O'Kane, C. J.; Floto, R. A.; Rubinsztein, D. C., *Nat. Chem. Biol.,* **2008,** *4*, 295-305.
6.   Martin, I.; Dawson, V. L.; Dawson, T. M., *Annu. Rev. Genomics Hum. Genet.* **2011,** *12*, 301-25.
7.   Hershko, A., *Angew. Chem. Int. Ed.,* **2005,** *44¸*5932-5943.
8.   Cook, W. J.; Jeffrey, L. C.; Carson, M.; Chen, Z.; Pickart, C. M., *J Biol Chem* **1992,** *267* (23), 16467-71.
9.   Pickart, C. M.; Fushman, D., *Current Opinion in Chemical Biology* **2004,** *8* (6), 610-616.
10.   Rahimi, N., *Molecular Cancer Therapeutics* **2012,** *11* (3), 538-548.
11.   Montalbetti, C. A. G. N.; Falque, V., *Tetrahedron* **2005,** *61* (46), 10827-10852.
12.   Hershko, A.; Ciechanover, A., *Annu Rev Biochem* **1998,** *67*, 425-79.
13.   Deshaies, R. J.; Joazeiro, C. A. P., RING domain E3 ubiquitin ligases. In *Annual Review of Biochemistry*, 2009; Vol. 78, pp 399-434.
14.   Schulman, B. A.; Wade Harper, J., *Nature Reviews Molecular Cell Biology* **2009,** *10* (5), 319-331.
15.   Dikic, I.; Wakatsuki, S.; Walters, K. J., *Nature Reviews Molecular Cell Biology* **2009,** *10* (10), 659-671.
16.   Woelk, T.; Sigismund, S.; Penengo, L.; Polo, S., *Cell Div* **2007,** *2*, 11.
17.   Hicke, L., *Nat Rev Mol Cell Biol* **2001,** *2* (3), 195-201.
18.   Bach, I.; Ostendorff, H. P., *Trends Biochem Sci* **2003,** *28* (4), 189-95.
19.   Varadan, R.; Assfalg, M.; Haririnia, A.; Raasi, S.; Pickart, C.; Fushman, D., *Journal of Biological Chemistry* **2004,** *279* (8), 7055-7063; Fushman, D.; Walker, O., *J Mol Biol* **2010,** *395* (4), 803-14.
20.   Hurley, James H.; Lee, S.; Prag, G., *Biochemical Journal* **2006,** *399* (Pt 3), 361-372.
21.   Coyne, E. S.; Wing, S. S., *F1000Res* **2016,** *5*.
22.   Jung, J. E.; Wollscheid, H. P.; Marquardt, A.; Manea, M.; Scheffner, M.; Przybylski, M., *Bioconjug Chem* **2009,** *20* (6), 1152-62.
23.   Shanmugham, A.; Fish, A.; Luna-Vargas, M. P.; Faesen, A. C.; El Oualid, F.; Sixma, T. K.; Ovaa, H., *J Am Chem Soc* **2010,** *132* (26), 8834-5.
24.   El Oualid, F.; Merkx, R.; Ekkebus, R.; Hameed, D. S.; Smit, J. J.; de Jong, A.; Hilkmann, H.; Sixma, T. K.; Ovaa, H., *Angewandte Chemie International Edition* **2010,** *49* (52), 10149-10153.

25.	Wilkinson, K. D.; Gan-Erdene, T.; Kolli, N., *Methods Enzymol* **2005,** *399*, 37-51.

26.	Kumar, K. S. A.; Haj-Yahya, M.; Olschewski, D.; Lashuel, H. A.; Brik, A., *Angewandte Chemie-International Edition* **2009,** *48* (43), 8090-8094.

27.	Haj-Yahya, M.; Ajish Kumar, K. S.; Erlich, L. A.; Brik, A., *Biopolymers* **2010,** *94* (4), 504-10.

28.	Virdee, S.; Kapadnis, P. B.; Elliott, T.; Lang, K.; Madrzak, J.; Nguyen, D. P.; Riechmann, L.; Chin, J. W., *J Am Chem Soc* **2011,** *133* (28), 10708-11.

29.	Yang, R.; Pasunooti, K. K.; Li, F.; Liu, X. W.; Liu, C. F., *J Am Chem Soc* **2009,** *131* (38), 13592-3.

30.	Marin, J.; Didierjean, C.; Aubry, A.; Casimir, J. R.; Briand, J. P.; Guichard, G., *J Org Chem* **2004,** *69* (1), 130-41.

31.	Dawson, P. E.; Kent, S. B., *Annu Rev Biochem* **2000,** *69*, 923-60.

32.	Johnson, E. C.; Kent, S. B., *J Am Chem Soc* **2006,** *128* (20), 6640-6.

33.	Mitra, N., Incorporating Unnatural Amino Acids into Recombinant Proteins in Living Cells. MATER METHODS, 2013; Vol. 3, p 204.

34.	Padrón, J. M.; Kokotos, G.; Martín, T.; Markidis, T.; Gibbons, W. A.; Martín, V. c. S., *Tetrahedron: Asymmetry* **1998,** *9* (19), 3381-3394.

35.	Fanning, K. N.; Sutherland, A., *Tetrahedron Letters* **2007,** *48* (48), 8479-8481.

36.	Gunatilaka, L. A. A.; Karunaratne, V.; Sandanayake, V. P.; Sotheeswaran, S., Conversion of Cyclic Enones into 1,3-Endodisulphides. A Synthesis of Carvone Endodisulphide. J. Chem. Research (S), 1990; pp 160-161; Tomisawa, K.; Kameo, K.; Matsunaga, T.; Saito, S.; Hosoda, K.; Asami, Y.; Sota, K., *Chem Pharm Bull (Tokyo)* **1985,** *33* (6), 2386-94; Chakor, N.; Dallavalle, S.; Musso, L.; Moretti, M., *Tetrahedron Letters* **2008,** *49* (34), 5056-5058; Condon, M. E.; Petrillo, E. W., Jr.; Ryono, D. E.; Reid, J. A.; Neubeck, R.; Puar, M.; Heikes, J. E.; Sabo, E. F.; Losee, K. A.; Cushman, D. W.; Ondetti, M. A., *J Med Chem* **1982,** *25* (3), 250-8; Zheng, T.-C.; Burkart, M.; Richardson, D. E., *Tetrahedron Letters* **1999,** *40* (4), 603-606.

37.	Robert, F.; Héritier, J.; Quiquerez, J.; Simian, H.; Blank, I., *Journal of Agricultural and Food Chemistry* **2004,** *52* (11), 3525-3529; Vermeulen, C.; Collin, S., *Journal of Agricultural and Food Chemistry* **2002,** *50* (20), 5654-5659; Martínez-Ramos, F.; Vargas-Díaz, M. a. E.; Chacón-García, L.; Tamariz, J. n.; Joseph-Nathan, P.; Zepeda, L. G., *Tetrahedron: Asymmetry* **2001,** *12* (22), 3095-3103; Burkhard, J. A.; Guérot, C.; Knust, H.; Rogers-Evans, M.; Carreira, E. M., *Organic Letters* **2010,** *12* (9), 1944-1947.

38.	Nair, D. P.; Podgórski, M.; Chatani, S.; Gong, T.; Xi, W.; Fenoli, C. R.; Bowman, C. N., *Chemistry of Materials* **2014,** *26* (1), 724-744.

39.	Yu, L.; Wang, L.-H.; Hu, Z.-T.; You, Y.-Z.; Wu, D.-C.; Hong, C.-Y., *Polymer Chemistry* **2015,** *6* (9), 1527-1532; Griesbaum, K.; Oswald, A. A.; Quiram, E. R.; Naegele, W., *The Journal of Organic Chemistry* **1963,** *28* (8), 1952-1957; Biermann, U.; Butte, W.; Koch, R.; Fokou, P. A.; Turunc, O.; Meier, M. A.; Metzger, J. O., *Chemistry* **2012,** *18* (26), 8201-7; Mueller, W. H., *The Journal of Organic Chemistry* **1966,** *31* (10), 3075-3079; Connal, L. A.; Kinnane, C. R.; Zelikin, A. N.; Caruso, F., *Chemistry of Materials* **2009,** *21* (4), 576-578.

40.	Prucker, O.; Rühe, J., *Macromolecules* **1998,** *31* (3), 602-613.

41.   Zeynizadeh, B.; Behyar, T., *Bulletin of the Chemical Society of Japan* **2005,** *78* (2), 307-315.

42.   Gololobov, Y. G.; Zhmurova, I. N.; Kasukhin, L. F., *Tetrahedron* **1981,** *37* (3), 437-472.

43.   Abdel-Magid, A. F.; Carson, K. G.; Harris, B. D.; Maryanoff, C. A.; Shah, R. D., *The Journal of Organic Chemistry* **1996,** *61* (11), 3849-3862.

44.   Yang, R.; Pasunooti, K. K.; Li, F.; Liu, X. W.; Liu, C. F., *Chem Commun (Camb)* **2010,** *46* (38), 7199-201.

45.   Sun, Y.-Q.; Chen, M.; Liu, J.; Lv, X.; Li, J.-f.; Guo, W., *Chemical Communications* **2011,** *47* (39), 11029-11031; Kimmel, K. L.; Robak, M. T.; Ellman, J. A., *Journal of the American Chemical Society* **2009,** *131* (25), 8754-8755; Kobayashi, N.; Iwai, K., *The Journal of Organic Chemistry* **1981,** *46* (9), 1823-1828; Li, H.; Wang, J.; Zu, L.; Wang, W., *Tetrahedron Letters* **2006,** *47* (15), 2585-2589.

46.   Liu, J.; Castaneda, C. A.; Wilkins, B. J.; Fushman, D.; Cropp, T. A., *Bioorg Med Chem Lett* **2010,** *20* (19), 5613-6.

47.   Ai, H.-w.; Lee, J. W.; Schultz, P. G., *Chemical Communications* **2010,** *46* (30), 5506-5508.

48.   Kang, J.; Macmillan, D., *Org Biomol Chem* **2010,** *8* (9), 1993-2002.

49.   McGinty, R. K.; Kohn, M.; Chatterjee, C.; Chiang, K. P.; Pratt, M. R.; Muir, T. W., *ACS Chem Biol* **2009,** *4* (11), 958-68.

50.   Castaneda, C. A.; Liu, J.; Kashyap, T. R.; Singh, R. K.; Fushman, D.; Cropp, T. A., *Chem Commun (Camb)* **2011,** *47* (7), 2026-8.

51.   Baldwin, J. E., *Journal of the Chemical Society, Chemical Communications* **1976,** (18), 734-736.

52.   Yu, H.-A.; Kim, S.-G.; Kim, E.-J.; Lee, W.-J.; Kim, D.-O.; Park, K.; Park, Y.-C.; Seo, J.-H., *Protein Expression and Purification* **2007,** *56* (1), 20-26.

53.   Petersson, E. J.; Goldberg, J. M.; Wissner, R. F., *Phys Chem Chem Phys* **2014,** *16* (15), 6827-37; McGouran, J. F.; Kramer, H. B.; Mackeen, M. M.; di Gleria, K.; Altun, M.; Kessler, B. M., *Org Biomol Chem* **2012,** *10* (17), 3379-83.

54.   Kosen, P. A., *Methods Enzymol* **1989,** *177*, 86-121.

55.   Ranjani, V.; Assfalg, M.; Fushman, D., Using NMR Spectroscopy to Monitor Ubiquitin Chain Conformation and Interactions with Ubiquitin-Binding Domains. In *Methods in Enzymology*, Academic Press: 2005; Vol. Volume 399, pp 177-192.

56.   Pickart, C. M.; Raasi, S., *Methods Enzymol* **2005,** *399*, 21-36.

57.   Virdee, S.; Ye, Y.; Nguyen, D. P.; Komander, D.; Chin, J. W., *Nat Chem Biol* **2010,** *6* (10), 750-7.

58.   Castañeda, C.; Liu, J.; Chaturvedi, A.; Nowicka, U.; Cropp, T. A.; Fushman, D., *Journal of the American Chemical Society* **2011,** *133* (44), 17855-17868.

59.   Isidro-Llobet, A.; Álvarez, M.; Albericio, F., *Chemical Reviews* **2009,** *109* (6), 2455-2504.

60.   Aimoto, S., *Biopolymers* **1999,** *51* (4), 247-65.

61.   Mukai, T.; Kobayashi, T.; Hino, N.; Yanagisawa, T.; Sakamoto, K.; Yokoyama, S., *Biochem Biophys Res Commun* **2008,** *371* (4), 818-22; Kobayashi, T.; Yanagisawa, T.; Sakamoto, K.; Yokoyama, S., *J Mol Biol* **2009,** *385* (5), 1352-60.

62.   Streu, C.; Meggers, E., *Angewandte Chemie International Edition* **2006,** *45* (34), 5645-5648.

63.   Trost, B. M.; Dong, L.; Schroeder, G. M., *Journal of the American Chemical Society* **2005,** *127* (29), 10259-10268; Antos, J. M.; Francis, M. B., *Current Opinion in Chemical Biology* **2006,** *10* (3), 253-262.

64.   Ryabov, Y.; Fushman, D., *Proteins* **2006,** *63* (4), 787-96.

65.   Weikart, N. D.; Sommer, S.; Mootz, H. D., *Chem Commun (Camb)* **2012,** *48* (2), 296-8.

66.   Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G., *Journal of the American Chemical Society* **2002,** *124* (31), 9026-9027.

67.   Trempe, J. F.; Brown, N. R.; Lowe, E. D.; Gordon, C.; Campbell, I. D.; Noble, M. E.; Endicott, J. A., *Embo j* **2005,** *24* (18), 3178-89.

68.   Iphofer, A.; Kummer, A.; Nimtz, M.; Ritter, A.; Arnold, T.; Frank, R.; van den Heuvel, J.; Kessler, B. M.; Jansch, L.; Franke, R., *Chembiochem* **2012,** *13* (10), 1416-20.

69.   Yin, L.; Krantz, B.; Russell, N. S.; Deshpande, S.; Wilkinson, K. D., *Biochemistry* **2000,** *39* (32), 10001-10.

70.   Valkevich, E. M.; Guenette, R. G.; Sanchez, N. A.; Chen, Y. C.; Ge, Y.; Strieter, E. R., *J Am Chem Soc* **2012,** *134* (16), 6916-9.

71.   Trang, V. H.; Valkevich, E. M.; Minami, S.; Chen, Y. C.; Ge, Y.; Strieter, E. R., *Angew Chem Int Ed Engl* **2012,** *51* (52), 13085-8.

72.   Chen, J.; Ai, Y.; Wang, J.; Haracska, L.; Zhuang, Z., *Nat Chem Biol* **2010,** *6* (4), 270-2.

73.   Chatterjee, C.; McGinty, R. K.; Fierz, B.; Muir, T. W., *Nat Chem Biol* **2010,** *6* (4), 267-9.

74.   Yang, K.; Gong, P.; Gokhale, P.; Zhuang, Z., *ACS Chem Biol* **2014,** *9* (8), 1685-91.

75.   McGinty, R. K.; Kim, J.; Chatterjee, C.; Roeder, R. G.; Muir, T. W., *Nature* **2008,** *453* (7196), 812-6.

76.   Kosower, N. S.; Kosower, E. M., *Methods Enzymol* **1995,** *251*, 123-33.

77.   Chen, J.; Ai, Y.; Wang, J.; Haracska, L.; Zhuang, Z., *Nat Chem Biol* **2010,** *6* (4), 270-272.

78.   Hemelaar, J.; Borodovsky, A.; Kessler, B. M.; Reverter, D.; Cook, J.; Kolli, N.; Gan-Erdene, T.; Wilkinson, K. D.; Gill, G.; Lima, C. D.; Ploegh, H. L.; Ovaa, H., *Mol Cell Biol* **2004,** *24* (1), 84-95; Reyes-Turcu, F. E.; Horton, J. R.; Mullally, J. E.; Heroux, A.; Cheng, X.; Wilkinson, K. D., *Cell* **2006,** *124* (6), 1197-208.

79.   Amerik, A. Y.; Hochstrasser, M., *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **2004,** *1695* (1–3), 189-207.

80.   Schanda, P.; Brutscher, B., *Journal of the American Chemical Society* **2005,** *127* (22), 8014-8015.

81.   Varadan, R.; Walker, O.; Pickart, C.; Fushman, D., *J Mol Biol* **2002,** *324* (4), 637-47.

82.   Lai, M. Y.; Zhang, D.; Laronde-Leblanc, N.; Fushman, D., *Biochim Biophys Acta* **2012,** *1823* (11), 2046-56.

83.   Zhang, C.; Saunders, A. J., *Discov Med* **2009,** *8* (40), 18-22.

84.   Riedinger, C.; Boehringer, J.; Trempe, J.-F.; Lowe, E. D.; Brown, N. R.; Gehring, K.; Noble, M. E. M.; Gordon, C.; Endicott, J. A., *Journal of Biological Chemistry* **2010,** *285* (44), 33992-34003.

85.   Varadan, R.; Assfalg, M.; Raasi, S.; Pickart, C.; Fushman, D., *Mol Cell* **2005,** *18* (6), 687-98; Mueller, T. D.; Kamionka, M.; Feigon, J., *J Biol Chem* **2004,** *279* (12), 11926-36.

86.   Zhang, D.; Raasi, S.; Fushman, D., *J Mol Biol* **2008,** *377* (1), 162-80.

87.   Zhang, D.; Chen, T.; Ziv, I.; Rosenzweig, R.; Matiuhin, Y.; Bronner, V.; Glickman, M. H.; Fushman, D., *Mol Cell* **2009,** *36* (6), 1018-33.

88.   Edelmann, M. J.; Nicholson, B.; Kessler, B. M., *Expert Rev Mol Med* **2011,** *13*, e35.

89.   Bedford, L.; Lowe, J.; Dick, L. R.; Mayer, R. J.; Brownell, J. E., *Nat Rev Drug Discov* **2011,** *10* (1), 29-46.

90.   Basle, E.; Joubert, N.; Pucheault, M., *Chem Biol* **2010,** *17* (3), 213-27.

91.   Piotrowski, J.; Beal, R.; Hoffman, L.; Wilkinson, K. D.; Cohen, R. E.; Pickart, C. M., *J Biol Chem* **1997,** *272* (38), 23712-21.